Behavioural changes of a long-ranging diver in response to oceanographic conditions

Theoni Photopoulos

BL5019 Research Project
MRes Environmental Biology

September 2007

Behavioural changes of a long-ranging diver in response to oceanographic conditions

Student: Theoni Photopoulos

Supervisors: Martin Biuw
Sea Mammal Research Unit
Gatty Marine Laboratory
University of St. Andrews
St. Andrews
Scotland, KY16 8LB

Len Thomas
Center for Research into Ecological and Environmental Modelling
University of St Andrews
The Observatory
St Andrews
Scotland, KY16 9LZ

Abstract

The development of an animal-borne instrument that can record oceanographic measurements (CTD-SRDL) has enabled the collection of oceanographic data at a scale relevant to the counterpart behavioural data, both in time and 3-dimensional space. This has advanced the potential for studies of the behaviour of deep-diving marine animals and the way in which they respond to their environment, yet the nature of the data delivered by CTD-SRDLs presents substantial analytical challenges and places constraints on its biological interpretation. Behavioural and environmental data, collected using CTD-SRDLs deployed on southern elephant seals (*Mirounga leonina*) from the South Georgia subpopulation in 2004 and 2005, are analysed for 13 females and 4 males (21,015 dives). Compressed dive profiles are used to classify individual dives into six distinct types based on their 2-dimensional time-depth characteristics using random forest classification. The relationship between dive type and environmental variables, derived from oceanographic data recorded on board the animals, is investigated in the context of regression analysis, employing a multinomial model, as well as independently fitted Generalized Linear Models (GLM) and Generalized Additive Models (GAM) for each dive type. Regression is not found to be an appropriate method for analysing abstracted behavioural dive data, and other methods are suggested. We show that functional specializations can be manifested within a dive type, using square bottom dives (SQ) as an example. The usefulness of dive classification is discussed in the context of behavioural interpretation, and validity of the ecological functions attached to each class. Preliminary analyses are important drivers of further research into improving the interpretability of abstracted behavioural data, and developing efficient, standardized methods for widespread application to this type of data, which is obtained in abundance via satellite telemetry.

Acknowledgements

Many thanks to my supervisors, Martin Biuw and Len Thomas, for all their time, enthusiasm, and guidance during the course this project. Thank you to SMRU and Mike Fedak for access to this dataset. Special thanks go to Monique MacKenzie for valuable help and suggestions during analysis. I would also like to thank Lars Boehme for help with the oceanographic side of things, Clint Blight for much technical support, and Jason Matthiopoulos for his time and ideas. Last but not least, thank you to my parents for funding my post graduate degree, as well as my sister Christie and Simon Elwen for support throughout this year.

Table of Contents

1. Introduction

1.1. Studying Behaviour in Diving Marine Mammals

Marine mammals are a particularly challenging group of animals in which to study behaviour because neither their behaviour nor their environment is easily observed. Many marine mammals spend most if not all of their lives at sea, often spending extended periods at depth. Interpreting the behaviour of deep-diving animals requires information about the marine environment at relevant scales in time and 3-dimensional space. Remote sensing datasets that are used extensively to provide an environmental context for animal behaviour often only provide information on sea surface characteristics and this constrains interpretation of behavioural data in an appropriate environmental framework. Technological developments in the use of satellite telemetry for wildlife studies, based on the Argos satellite system[1], as well as development of the instruments themselves have opened up the field of bio-logging to more in-depth studies of animal behaviour in relation to their environment (Fedak *et al.* 2002, Fedak 2003). The development of one particular animal-borne instrument that can record oceanographic measurements, the Conductivity-Temperature-Depth Satellite Relay Data Logger (CTD-SRDL) (Lydersen *et al.* 2002), has enabled the collection of high-resolution oceanographic data from inaccessible and data deficient habitats, and advanced the potential for behavioural studies of deep-diving marine animals (Biuw *et al.* 2007, Boehme *et al.* in press, Bailleul *et al.* 2007, Fedak 2003, Charassin *et al.* 2002, Boehlert *et al.* 2001). Southern elephant seals (*Mirounga leonina*) are well suited to studies at the interface of animal behaviour and the physical marine environment since they are numerous, deep-diving and wide ranging throughout the Southern Ocean. These features allow them to exploit prey resources over a range of spatial and temporal scales, and are likely to make them sensitive to large-scale changes in their environment. In this study, behavioural and environmental data collected using CTD-SRDLs deployed on southern elephant seals from the South Georgia subpopulation in 2004 and 2005, are analysed for 13 females and 4 males (21,015 dives).

1.2. A Special Case of Behavioural Data

All data recorded on board CTD-SRDLs are transmitted via the Argos satellite system to ground stations, a process that is only possible while the animals are at the surface. Three factors: the short surfacing intervals characteristic of elephant seals (~2min, Le Boeuf *et al.* 2000b), the narrow bandwidth available via the Argos satellite system (Fedak *et al.* 2002), together with the desire to maximise the amount of information delivered, by maximising the number of transmissions per day and thus prolonging battery life; make it necessary for both behavioural and environmental data to be compressed before they are relayed to the Argos ground station.

[1] Argos is a worldwide location and data collection system dedicated to the environmental research, which comprises three elements: (1) platforms that hold Argos transmitters (here CTD-SRDLs), (2) satellites that receive the transmitters' signals and (3) ground stations that receive the messages transmitted via satellite from the platforms (Argos 2007).

Consequently, the 2-dimensional (time vs. depth) dive profiles delivered by the instrument consist of only 4 points; the 4 points of maximum inflection in the time-depth trace of the dive (Fedak *et al.* 2001, Fedak *et al.* 2002). Information is necessarily lost in the compression process but inflection-point dive records have nevertheless been shown to perform well when compared against Time-Depth Recorder (TDR) data for leatherback turtles (Myers *et al.* 2006), where depth is measured every 10 or 30 seconds. These compressed profiles are therefore sufficient to classify individual dives into distinct types based on their 2-dimensional time-depth characteristics. Six dive classes are used for the purpose of this analysis including square bottom dives (SQ), U-shaped dives (U), V-shaped dives (V), wiggle dives (W), drift dives (DR) and root shaped dives (R) (Fig. 1-6 A), based on previous work on dive function and classification in elephant seals and other pinnipeds (Hindell *et al.* 1991, Davis *et al.* 2003). Ecological functions that have been loosely attached to these classes are “benthic foraging” for SQ dives, “travelling” for U and V dives, “pelagic foraging” for W dives, “resting” for DR dives and “exploration” in R dives. Classification is carried out using a random forest tree-building classification method based on the algorithm by Breiman (2001), because the amount of data made manual classification impractical. This method has been developed specifically to deal with inflection-point dive data such as those delivered by CTD-SRDL tags, but is flexible enough so that it can also be used to classify continuous data of the type obtained from TDRs (Biuw *et al.* in prep.). The practical advantage of allocating types to dive data is that it produces a response variable that can then be coupled with environmental data. It can also be a useful approach when faced with abstracted data, such as compressed dive profiles, and when trying to detect changes in diving behaviour. However, assigning classes to animal behaviour based on observed parameters should be exercised with caution as animals cannot be expected to behave in a categorised fashion and there are almost certainly gradations between dive types. Furthermore, mistaken labelling of behaviours can lead to masking of the true patterns in the data or generating artificial ones.

The commercial availability of satellite-linked data recorders has lead to hundreds of loggers (Myers *et al.* 2006) being deployed on wildlife worldwide, including southern (Biuw *et al.* 2007) and northern elephant seals (Crocker *et al.* 2006), leopard seals (Kuhn *et al.* 2005), ringed seals (Lydersen *et al.* 2004), Mediterranean monk seals (Dendrinos *et al.* 2007), crabeater seals (Wall *et al.* 2007), leatherback turtles (James *et al.* 2006), olive ridley turtles (McMahon *et al.* 2007), narwhals (Lydersen *et al.* 2007), and beluga whales (Lydersen *et al.* 2002) to collect environmental and behavioural data. The environmental data recorded by animal-borne CTD-SRDLs, though constrained in resolution by the narrow bandwidth of the Argos satellite system, are well matched with the questions asked regarding the marine environment, since the marine environment can be characterised using the parameters recorded by the tag, namely temperature, pressure and salinity (derived from pressure, temperature and conductivity measurements). As a result, these environmental data are readily interpreted and utilised by oceanographers (Boehme *et al.* in press). In contrast, the data required to answer the behavioural questions being asked; such as, what conditions are different behaviours associated with, cannot as yet be obtained, insofar as elephant seal behaviour such as foraging strategies or feeding events cannot be directly observed or recorded using CTD-SRDLs. The data that are recorded, in this case inflection-point dive profiles, are used as surrogates under the assumption that they provide information about different animal behaviours, which renders biological interpretation challenging. The development of methods for analysing highly compressed behavioural and

oceanographic datasets independently is currently underway, but very few attempts have been made to relate the two in a coherent way (Bailleul *et al.* 2007, Biuw *et al.* 2007).

Several issues arise when dealing with the behavioural, inflection-point dive data obtained from SRDLs that need to be addressed in any analytical approach. These include non-independence between dives performed by the same animal, non-independence between consecutive dives, non-linearity of the response with respect to the covariates, and the effect of classification error. The issues that are touched upon within the scope of this study are non-linearity of the response and non-independence between dives within individuals. Suggestions are made as to how to address the remaining issues. An additional issue that relates to the nature of satellite telemetry is that data are not relayed for every single dive, which presents a problem if data are to be dealt with as a time series.

1.3. Analytical Routes and Challenges

There are at least three ways in which an analysis of this kind of abstracted behavioural information can be approached; (1) look at the raw data, exploratory data analysis, (2) formulate hypotheses based on the observed data and test them statistically, and (3) use a Markovian type analysis.

The most common first line of enquiry into this type of information is to (1) look at the raw data. An exploratory analysis serves to help identify patterns in the data, which can be quantified using statistical models. Goodness-of-fit measures can then quantify how well the model fits the data; in other words, how well it has captured the relationship between the response and its covariates.

A second way to approach the problem of understanding the relationship between animal behaviour and environmental data is to (2) formulate hypotheses based on the observed data and test them statistically. There are a number of reasons that this approach might not be suitable for highly derived data such as telemetry data. These stem from the fact that end product data are noisy as a result of errors being compounded over various stages of data processing. For example, it is not always possible to obtain geographical locations associated with individual dives due to the requirements of the uplink process employed by the Argos satellite system. Instead, dive locations are interpolated based on dive start time in relation to the time of the closest position fixes for adjacent dives, as is the case with bathymetric datasets where only some depths are sounded and the surface of the seabed interpolated from them. As a result, an analysis of dive type versus bathymetry already includes at least three separate sources of error, illustrating the difficulty in carrying out robust analyses on highly derived data.

A third route to solving the problem at hand is (3) to use a Markovian type analysis. In this case the time-series of dives representing the different types would be regarded as a discrete-time stochastic process where the distribution of the response is conditional on the history of the process and each observation is determined by the observation preceding it. Hidden Markov Models (HMM) extend this type of analysis by allowing for measurement errors in observing

the response. HMMs assume two stochastic processes: the underlying unobserved process and the observed process. The observed process is regarded as a distortion of the underlying process due to noise, and inferences regarding the underlying process are drawn from the observed process (Cappé *et al.* 2005). It is easy to see how a Markovian approach might make biological sense if dive types are not independent of each other, and also how HMM might be useful if dive classes are assumed to be surrogates for the underlying unobserved pattern in diving behaviour.

For the purpose of this thesis only the first of these approaches is attempted. A range of regression analyses of dive type are used to combine dive type with contemporaneous environmental data in an attempt to gain insight into the way in which behaviour relates to *in situ* environmental conditions. Diurnal, seasonal and sex differences in the response to environmental conditions were not investigated here but warrant inclusion in further analyses of dive type as a behavioural response (Bennet *et al.* 2001, Jonker & Bester 1998, Campagna *et al.* 1999).

1.4. Southern Elephant Seal Diving Behaviour at South Georgia

This work is motivated by the fact that satellite telemetry studies have provided good evidence as to where elephant seals go in the Southern Ocean (McConnell *et al.* 2002, Hindell *et al.* 2003), but little is known about how the environmental conditions they encounter influence their decision making process. The objective of this study is to determine whether individual dive types are correlated with the environmental variables derived from oceanographic data recorded onboard the animals. The relationships between the response and its covariates are examined here using a series of methods, starting from an exploratory data analysis, moving through three regression analyses: a Multinomial model, independently fitted Generalized Linear Models (GLM) and Generalized Additive Models (GAM) for each dive type. For the purpose of this analysis we assume that the classification of dives into types based on time-depth parameters is correct and that individual types correspond to distinct ecological functions. The validity of this assumption is discussed. In addition, the relationship between maximum dive depth and bathymetry in SQ dives is investigated for dives in water >1000m depth to determine whether animals dive to the seabed in areas where it is accessible to them.

A multinomial model was used to model the relationship between dive type and environmental covariates, as the response variable in question, dive type, is a multinomial one. To circumvent the use of the relatively inflexible multinomial model, separate regression models were fitted to binary data for each nominal response, here being each dive type, although Agresti (1996) warns this is a sub-optimal approach. The advantage of binary regressions is that modelling each dive type separately makes it possible to explore whether different covariates were important in predicting different dive types. This is the rationale behind fitting individual GLMs to binary data for individual dive types. A series of GAMs are also fitted to investigate whether the responses are non-linear with respect to the explanatory variables. The methods used to investigate maximum depth are Generalised Least Squares (GLS) and a Generalised Additive Mixed Model (GAMM). GLS is a mixed effects model that allows the specification of an error

structure but assumes a linear relationship between the response and covariate(s), whereas a GAMM is a more flexible version of a GLS model, within which it is possible to fit covariates as random effects and assume a non-linear relationship between the response and covariate(s). Each method and its results are examined in sequence and summarised in the discussion.

2. Methods

2.1. Data Collection

Dive profiles (21,015) and oceanographic data (8,250 CTD profiles) were recorded by animal-borne CTD-SRDLs for 17 adult southern elephant seals (13 females and 4 males) instrumented at South Georgia in January and February 2004 (6 females) and 2005 (4 males, 7 females) after the annual molt (Table 1 A). The instruments were designed and manufactured by the Sea Mammal Research Unit (St. Andrews, U.K.), incorporating a CTD sensor built by Valeport Ltd. (Totnes, U.K.). Tracks lasted on average two and a half months (mean ± SD, 73.3 ± 64.2 days), and ranged from 17 to 207 days in length. This is shorter than in previous studies (Biuw *et al.* 2007) due to a software failure in the instruments deployed in 2005 (2005: mean ± SD, 46.8 ± 31.2, range 26 to 114; 2004: mean ± SD, 122 ± 83.5, range 17 to 207). As a result the useable sections of most of the tracks from 2005 animals do not start from South Georgia.

Behavioural and environmental data are collected by the instruments at high frequency and then compressed according to Fedak *et al.* (2002) before being transmitted, in order to maximise the amount of representative information transmitted via the narrow bandwidth of the Argos satellite relay system. Behavioural data recorded by CTD-SRDLs include the time-depth dive profile from which a suite of dive parameters can be derived (Table 1 B, where capital B indexes Appendix B). Dive profiles are summarized as 2-dimensional time-depth profiles consisting of four points selected as the points of maximum inflection where the dive trajectory changed most, which are computed on board the tag before transmission (Biuw *et al.* 2007, Fedak *et al.* 2002). The physical data delivered by the tag are temperature, pressure and conductivity summarised as 20-point profiles; for a detailed account of the quality control of these measurements see Biuw *et al.* (2007) and Boehme *et al.* (in press). The power requirement of measuring conductivity on the tag dictates that not every dive has a corresponding CTD profile. Instead, the aim is to obtain at least one CTD profile per 6hour period that will represent the deepest dives. Recording of a CTD profile is designed to commence on the ascent for dives that are deeper than a predetermined depth threshold (1000m depth). To ensure that at least one CTD profile is recorded for every 6hour period, and that each profile represents the deepest possible dive during that period, the depth threshold for recording CTD data is gradually reduced through the six hour period down to a minimum of approximately 300m depth, to increase the likelihood of a dive being deep enough to trigger the collection of CTD data. The resulting ratio of CTD profiles to dive profiles is one to four.

2.2. Data processing

All data processing and analysis is carried out in the R software environment (R Development Core Team, 2007). Dives included in the analysis are >6m in depth, have four non-identical inflection points and no missing values for dive parameters. The potential covariates used to predict dive type in this analysis are water mass and frontal zone in the Southern Ocean,

distance from mixed layer depth (MLD), distance from isothermal depth (ILD), temperature, salinity, maximum dive depth, bathymetry and an indicator variable to specify whether dives occurred in continental shelf waters (1) or in pelagic waters (0) and individual reference ID. Because CTD profiles are not available for every single dive, temperature and salinity values are ascribed to dives by using a weighted running average. Here, the weights are based on a standardized Gaussian distribution over a 7-day period, where the mean of the distribution falls at the time of the dive of interest, and successively lower weights are ascribed to CTD profiles at greater time periods before or after this time. This makes it possible to extract water mass and frontal zone for those locations. Missing data for MLD, ILD, temperature at 200m and 500m are also interpolated in a similar way over a 5-day bandwidth. The trade-off involved with choosing a bandwidth for interpolation is that a very narrow bandwidth gives high accuracy in the interpolated data but leads to the loss of a large proportion of the dive data that is not paired with environmental data. Conversely, a wide bandwidth gives low accuracy but allows a large proportion of the dive data to be paired with environmental data. For the purpose of this analysis the latter approach is chosen to retain a large dataset.

Water mass is derived from temperature and salinity at maximum dive depth for all dive types using cut off values from Boehme *et al.* (in press) to define four water masses: surface water (SUR), Antarctic intermediate water (AAIW), upper circumpolar deep water (UCDW), deep water (DEEP). Frontal zone is derived by extracting interpolated temperature values at 200m and 500m depth using cut off values from Heywood & King (2002) to define four frontal zones: Subantarctic zone (SAZ), polar frontal zone (PFZ), Antarctic zone (AAZ), south of the Antarctic circumpolar current (S.ACC). Mixed layer depth, defined as the depth at which the vertical change in density exceeds a critical value, and isothermal layer depth, defined as the depth at which the vertical change in temperature exceeds a critical value are computed as outlined by Kara *et al.* 2000. Bathymetry is extracted from the two grid cell resolution ETOPO2v2g (NOAA 2007) dataset for all dive locations. The incorrect bathymetric data that are easily identifiable (i.e. on-land values) are associated with dives that occur very close to shore as a result of the steep relief of the South Georgia coastline. These values are replaced by resampling from the empirical distribution between 0 and 300m.

For R and W type dives, additional potential covariates are also water mass and distance from oceanographic features for an intermediate depth to assess whether it is more representative of the depth at which animals spend the most time while at depth. For W dives intermediate depth is taken as the median of the four points of highest inflection (i.e. the “wiggle” part of the dive) while for R dives mean depth is calculated for the plateau of the root shape of the dive.

2.3. Dive Classification

Compressed dives are classified into six distinct types: square (SQ), U-shaped (U), V-shaped (V), wiggle (W), drift (DR), root (R) (Fig. 1-6 A), according to depth-time parameters using the method described by Biuw *et al.* (in prep.). This method employs the random forest (RF) tree-building method implemented in the `randomForest` library in R (Liaw & Wiener 2002). A training dataset of approximately 10% of the dives was first classified manually and then used to

classify the remaining dives using the `randomForest` function in R (Breiman 2001). According to random forest classification each tree is built using a bootstrapped sample of the data so that trees are independent. For each node a random set of variables is selected at random with equal probability from a set of specified variables, of which the best are used for the split. The number of variables selected at each split is four and the number of trees grown is one thousand. The variables selected for the random forest classification are maximum dive depth, dive duration, bathymetry (ETOPO5), and fifteen dive parameters (Table 2 A). This bathymetry, already ascribed to dives in the database before analysis, was extracted from the ETOPO5 dataset, now outdated, which is why ETOPO2v2g was used in the analysis. The training dataset of manually classified dives consists of 3,000 dives classified manually based on visual cues, corresponding to 14.3% of dives used in the analysis (21,015 dives). The overall out of bag (OOB) error rate for this classification is 3.6% (Table 3 A). The error rates calculated by the random forest algorithm correspond to the predictions made at each bootstrap iteration for the data not in the sample using the tree grown by the bootstrap sample in question. The overall OOB error rate is then obtained by aggregating these error rates (Liaw & Wiener 2002). The `randomForest` output of predicted dive types is then used as input into the dive type analysis.

2.4. Dive Type Analysis

2.4.1. Exploratory Data Analysis

The proportion of different dive types and the relationships between dive type and potential covariates is first investigated graphically and using standard descriptive statistics. The relationship between dive type and each of the covariates, as well as the relationships between covariates are investigated using hypothesis tests. Pearson's Chi-squared tests for count data were used to test for significant differences in the distribution of dive types among levels of categorical variables (water mass, frontal zone, shelf[2]), and analysis of variance tests are used to test for significance in the relationship between dive type and continuous variables (distance from MLD, distance from ILD, bathymetry, maximum dive depth, temperature, salinity). The relationships between covariates are also tested using Pearson's Chi-squared tests for count data for categorical versus categorical variables, analysis of variance for continuous versus categorical variables and correlation for continuous versus continuous variables.

2.4.2. Multinomial Model

A multinomial model is fitted to the data using the `multinom` function in the `nnet` library for R (Venables & Ripley, 2002). Stepwise variable selection is carried out in both directions, starting with no covariates, using AIC. AIC is also used to decide between models.

[2] The continental shelf was defined as areas shallower than 700m depth based on the ETOPO2v2g dataset (2007)

2.4.3. Generalized Linear and Generalized Additive Models - GLMs and GAMs

GLMs are fitted to binary data, specifying a certain dive type or any of the others, for each of the dive types using `glm` (binomial family with a logit or probit link) in R (R Development Core Team, 2007). GAMs are fitted to the same binary data using `gam` (binomial family, using cubic regression splines for continuous variables) in R (Wood, 2006). Variable selection is carried out using automated forward and backward stepwise variable selection based on AIC for GLMs using `step`; where convergence problems arise, variable selection is carried out manually. Model selection for GLMs is carried out based on AIC and R squared values. Manual variable selection is carried out for GAMs and AIC is used for variable and model selection alike. The GAM fit to binary data for SQ dives is not possible due to lack of model convergence and it is therefore excluded from the results and discussion sections. The accuracy of predictions was assessed using plots of the confusion matrices of predicted versus observed occurrences of each dive type. The correlation structure in model residuals was investigated using correlograms of correlation versus lag in successive dives.

2.5. Maximum Dive Depth as a Response to Bathymetry in SQ dives

It has been suggested for several diving marine vertebrates that square bottom dives serve the function of bottom feeding dives (Charrassin *et al.* 2002, Hindell *et al.* 2003). To investigate this in southern elephant seals from South Georgia models are fitted to the relationship between maximum dive depth and bathymetry. Maximum dive depth in SQ dives is modelled using a Generalised Least Squares (GLS) model with a first order autocorrelation structure within individuals (Pinheiro *et al.* 2007), and also a Generalised Additive Mixed Model (GAMM) with individual animal fitted as a random effect, using a Gamma distribution so that the predictions are bounded by zero, and a first order correlation structure within individuals to investigate non-linearity in the response.

3. Results

3.1. Dive Type Analysis

3.1.1. Exploratory Data Analysis

General Trends

The most commonly performed dives across 17 individuals instrumented at South Georgia in 2004 and 2005, are U dives (mean 51.5%), followed by W dives (mean 29.7%), while the least commonly performed dives are SQ dives (average 0.8%) (Fig. 7 B). The highest proportion of dives occurs in Upper Circumpolar Deep Water (UCDW, 56.5%), and in the Antarctic zone (AAZ, 47.2%). Mean distance from maximum dive depth to the MLD and ILD is 411m (SD 199, range –287 to 1797) and 391m (SD 199, range –344 to 1743). Mean bathymetry across dives and individuals, corrected for values on land, is 3924m depth (SD 1367, range 1 to 8089) and mean maximum dive depth is 474m (SD 200, range 8 to 1902). Mean water temperature and salinity encountered at maximum dive depth are 2.43 degrees Celsius (SD 0.90, range -1.62 to 14.0) and 34.38 (SD 0.22, range 33.35 to 35.07). A detailed breakdown of parameter means by dive type is shown in Table 4 below. All relationships between the response and covariates but also among covariates are statistically significant at the $p < 0.001$ level. Water mass and distance from MLD and ILD for intermediate depths in R and W dives are not significantly different from the same variables calculated for maximum depth in those dive types and are therefore not considered further.

Table 4. Dive type frequency and mean (SD) values for continuous variables.

Dive Type	Freq. %	Distance from MLD	Distance from ILD	Temp. C	Salinity	Max. Dive Depth (m)	Bathymetry (m)	Prop. dives on shelf
SQ	0.67	152 (199)	141 (203)	2.89 (1.76)	34.15 (0.41)	213 (211)	1352 (1800)	0.70 (0.46)
U	51.51	443 (194)	426 (191)	2.36 (0.88)	34.40 (0.22)	597 (196)	3899 (1451)	0.07 (0.25)
V	1.44	623 (335)	611 (324)	2.50 (0.92)	34.35 (0.26)	687 (338)	3952 (1381)	0.03 (0.18)
W	29.73	350 (179)	327 (181)	2.55 (0.93)	34.32 (0.22)	416 (178)	4049 (1186)	0.03 (0.17)
DR	3.96	350 (152)	322 (151)	2.29 (0.93)	34.37 (0.21)	407 (153)	3724 (1371)	0.07 (0.26)
R	12.69	433 (203)	408 (202)	2.43 (0.82)	34.39 (0.82)	495 (203)	3924 (1235)	0.04 (0.19)

Bathymetry and the Continental Shelf

The bathymetry associated with all dive types except SQ dives has a roughly symmetric, unimodal distribution (Fig. 8 B). SQ dives seem to occur mostly over shallow water <1000m in depth, but also over much deeper water (4000-6000m) (Fig. 8 B). Of all six dive types SQ dives

are the only type that occurred more frequently in shelf waters than in pelagic waters (Fig. 9 B). Bathymetry and the proportion of dives made in shelf waters are statistically significantly different between different dive types ($F_5 = 123$, $p < 0.001$; Chi-squared = 1269, $p < 0.001$).

Frontal Zones and Water Masses

The most frequently visited water mass for all dive types, except SQ dives, is upper circumpolar deep water (UCDW), in which 56.6% of all dives occur (Fig. 10 B). For SQ dives, the most used water mass is surface water (SUR, 59.7%), which is probably explained by the comparatively shallow maximum dive depths for this dive type (mean 209m, SD 211, range 8 to 694). With the exception of SQ dives, all dive types follow the same overall pattern with the largest proportion of dives (47.2%) occurring in the Antarctic zone (AAZ) and the smallest (0.08%) south of the Antarctic circumpolar current (S.ACC) (Fig. 11 B). This is in contrast with data from females from the Macquarie Island population, which are reported to have spent 44.6% of the time in the polar frontal zone (PFZ) (Hindell *et al.* 2003). SQ dives are most frequently performed in the Antarctic and Subantarctic zones (AAZ, 58.9 % and SAZ, 31.8%) and least in the polar frontal zone and south of the Antarctic circumpolar current (PFZ, 0.06% and S.ACC, 0.03%, respectively) (Fig. 11 B). Water masses and frontal zones associated with different dive types are statistically significantly different (Chi-squared = 611, $p < 0.001$; Chi-squared = 407, $p < 0.001$).

Mixed and Isothermal Layer Depth (MLD and ILD)

The overall trend in distance from both MLD and ILD is for animals to dive through the mixed or isothermal layer (considered zero in both cases) to an average of 411m and 390m (mean). SQ dives seem to occur closest to both the MLD and ILD than other dive types (MLD: mean 148m, SD 199, range –287 to 677; ILD: mean 137m, SD 202, range –129 to 678) (Fig. 12, 13 B). It is possible that this is an artifact of animals diving down to the continental shelf, whose depth may often be similar to the MLD (mean ± SD, 64 ± 61, range 13 to 325m) or ILD (mean ± SD, 84 ± 47, range 15 to 482m). There was no compelling evidence for a close coupling between dive depth and MLD or ILD. Distance of maximum depth from MLD and ILD were both significantly different between different dive types ($F_5 = 353$, $p < 0.001$; $F_5 = 390$, $p < 0.001$).

Individual Effects

An unexplained software failure in the instruments deployed in 2005 rendered the majority of each track unusable for the 11 animals instrumented in that year (Table 1 A). As a result, the informative sections of tracks for that year are short excerpts of the full tracks, and often do not include the animals' departure from the island of South Georgia, where the SRDLs were deployed. The brevity of 2005 tracks prevents valid comparisons between individuals since the

information obtained from the 2005 tags is not representative of the range of conditions experienced by the instrumented animal. This is likely to have artificially enhanced the differences in dive type frequency between individuals and years and limited the scope for statistical inference based on the environmental conditions that were encountered.

In the data that were obtained, the relative occurrence of dive types varies on an individual basis, with some animals performing up to 10% SQ dives and other performing none at all (Fig. 14 B). There is also substantial variation in water masses and frontal zones used by individuals (Fig. 15, 16 B).

3.1.2. Multinomial Model

Variables selected in the final model by forward and backward stepwise selection based on AIC include water mass, frontal zone, temperature, salinity, distance from ILD, shelf, maximum dive depth and individual. Here, individual was included by the automated stepwise variable selection rather than purposefully retained as in the case of the GAMs. The multinomial model explained 30.7% of the deviance in the data. The fit of the model to data was poor based on a confusion matrix of observed versus predicted dive types as the model predicted mainly U and W dives, which are the two most commonly performed dive types, and failed to predict any SQ dives, which were the least common dive type (Table 5).

Table 5. Confusion matrix of predicted versus observed dive types for the multinomial model

	Observations					
Predictions	SQ	U	V	W	DR	R
SQ	0	0	0	0	0	0
U	54	9319	209	3778	587	1963
V	0	9	1	0	0	0
W	23	1522	31	2597	254	668
DR	0	0	0	0	0	0
R	0	0	0	0	0	0

3.1.3. Generalized Linear and Generalized Additive Models - GLMs and GAMs

During manual variable selection in the GAMs, individual is purposefully retained in the model to incorporate it as a fixed effect. This leads to many other covariates being retained in the model compared to automated variable selection in the GLMs, which excludes many more covariates (Table 6). This difference between the two types of models does not seem to have had large effects on the percentage of deviance explained by each (Table 7). Neither the linear nor the additive approach appeared to explain more than 22.8% of the deviance in any one dive type. Confusion matrices and plots of results for each model illustrate the overall poor fit to the data (Fig. 17-27 A, Table 8-18 A). Plotting the smooth functions on the scale of the response for the GAMs illustrates that the additive models appear to overfit to these data (Fig. 28, 30, 32, 34,

36 B). Plots on the scale of the linear predictor show the standard errors (Fig. 29, 31, 33, 35, 37 B). Correlograms for each of the GAMs suggested that the correlation in the model residuals is low (Fig. 38-42 B).

Table 6. Explanatory variables retained in the GLMs by AIC using manual and where possible automated variable selection, and the GAMs by AIC using manual variable selection.

Dive Type	Model	Wmass	Fzone	Temp	Sal	distMLD	distILD	Bathy	Shelf	Max. Dep	Individual
SQ	GLM	x		x		x			x		
	GAM	-	-	-	-	-	-	-	-	-	-
U	GLM	x	x				x		x	x	x
	GAM	x	x	x	x	x	x	x		x	x
V	GLM		x	x	x	x	x			x	
	GAM	x	x	x	x	x	x	x	x	x	x
W	GLM	x	x				x		x	x	x
	GAM	x	x	x	x	x	x	x	x	x	x
DR	GLM	x		x			x				
	GAM	x	x	x	x		x	x	x	x	x
R	GLM	x						x	x	x	x
	GAM	x		x	x		x	x		x	x

Table 7. Comparative table of model performance for GLM and GAM fits based on confusion matrices of predicted versus observed occurrence of each dive type using binary data. Values represent the proportion of failures (False positive: prediction 1, observation 0) and successes (True: prediction 1, observation 1) of the models to predict correctly.

	GLM predictions			GAM predictions		
Dive Type	% False Positive	% True	% Deviance explained	% False Positive	% True	% Deviance explained
SQ	15.1	75.3	22.8	-	-	-
U	37.5	63.9	6.3	36.3	65.0	8.0
V	0.4	2.9	12.5	0.3	3.7	21.0
W	18.9	44.3	8.0	18.3	45.6	10.0
DR	2.4	5.8	2.8	1.9	6.8	7.0
R	10.3	14.6	0.8	9.8	15.4	1.5

3.2. Maximum Dive Depth as a Response to Bathymetry in SQ dives

Both models fitted to the relationship between maximum dive depth in SQ dives and bathymetry in water shallower than 1000m show a significant relationship between the two, and give similar predictions (Fig. 43-48 B). The GLS model explains 91.5% of the variance in the data (R sq. value) but predicts implausible values for maximum depth (Fig. 47 B). The GAMM explains 85.4% of the variance in the data (R sq. value). The smooth function on the scale of the linear predictor suggests that the relationship between maximum dive depth and bathymetry in SQ dives within <1,000m is largely linear (Fig. 49 B). Despite substantial variation between individuals, at least some animals appear to dive to the seabed when in continental shelf waters.

4. Discussion

4.1. Challenges of Abstracted Behavioural Data Obtained via Telemetry

The challenges that are involved in the interpretation of behavioural dive data obtained from SRDLs have so far impeded their meaningful analysis in the context of the counterpart environmental data. The non-independence between dives performed by the same animal, non-independence between consecutive dives, non-linearity of the response with respect to the covariates, and the effect of dive classification error, impose complex rules upon an already complicated problem. Below are outlined some methods available for dealing with these challenges in a regression framework.

Within the binary regression framework the issue of non-independence dives performed by the same animal can be addressed in two ways, (1) individual can be fitted as a fixed effect which results in 17 different factors levels, or (2) individual can be fitted as a random effect in which case the model assumes the same mean response for each animal but allows the intercept to vary. Mixed effects models give average values of the response variable at the population level. The argument for using mixed effects models in this analysis would be that the dataset includes multiple repeat measurements from 17 individuals. The requirements of a mixed model include a) having a large enough sample size, and b) that the animals are sampled randomly with respect to the population of interest. Given that only 17 animals are represented in this dataset and that those animals were captured opportunistically for instrumentation, neither of these assumptions holds for the data at hand, therefore mixed effects models are not suitable and hence not employed in this analysis. In addition, it is not practically meaningful to compare effects across individuals for which there is substantial variation in track length. For an informative comparison between animals, to further investigate individual variation, tracks of approximately equal length should be chosen, during which the animals have sampled the same range of environmental conditions. Dive type could then be compared for the same set of environmental conditions. Regarding the non-independence of consecutive dives, this is addressed in the current study by means of generating correlograms of the Pearson's residuals for GAMs. If the models predict the occurrence of subsequent dives of the same type, this will in theory result in a correlation structure in the model residuals for successive dives within individuals. However, correlograms for the fitted GAMs showed a very weak correlation structure in model residuals.

Errors in RF classification can arise in two ways, firstly poor performance of the human classifier during manual classification of the training dataset, and secondly poor performance of the RF algorithm itself. Which of the two, if either, is manifesting itself in the results can be shown by calibrating manual classification by only using dives whose type has been agreed upon by multiple human classifiers, and refitting models using only the manually classified training set. If the dive classes themselves are incorrect, a reworking of the functionally distinct dive types will be required. The lack of functionally meaningful dive classes to choose from when assigning a given dive to a class will undoubtedly confound any patterns in the relationship between dive type and environmental variables. A way to address this issue would be to repeat the binary regression using a subset of the data containing only dive types for which

a functional basis has been best documented in elephant seals. Conversely, should poor performance of the RF algorithm be the reason behind poor model fit, refitting models to the manually classified random subset of dives used as the training dataset for the RF classification, will in theory result in improved predictive power. Lastly, if the absolute classification of dives into types is what is masking trends, incorporating the uncertainty associated with each dive belonging to a given type may improve the fit if probabilities are commonly very similar for several types. This could be implemented by using the probabilities returned by RF as weights in a binary regression.

4.2. Current Dive Type Analysis

The regression approach employed in this study in an attempt to explore the relationship between dive type and environmental characteristics in elephant seals gives inconclusive results. Due to the poor predictive power of the models used, results were not regarded as biologically significant and therefore not interpreted. The only striking pattern identified in the raw data during exploratory analysis was the relationship between maximum dive depth in SQ dives and bathymetry, which was analysed separately. The multinomial model explained a reasonable proportion of variability but this was limited only to the most common dive types. There was slight improvement in model fit from the linear to the additive counterpart models for each dive type, however the small proportion of variability explained by these models suggests that neither a linear nor additive approach yields models with compelling fits to dive type in southern elephant seals. The overall emergent trends in the data are in agreement with previous studies on southern elephant seals, namely that they spend most of the time at sea diving in UCDW, and in the AAZ and PFZ. This is in agreement with data from animals instrumented at the Kerguelen Islands, according to which the polar front is one of the two regions where animals displayed increased foraging activity and body condition, the other being the Antarctic continental shelf (Bailleul *et al.* 2007). Hindell *et al.* (2003) found that animals from the Macquarie Island population spent almost half of the time at sea in the PFZ.

If we assume that the classification of dives into types, based on dive shape parameters, accurately represents functionally different behaviours in elephant seals, predicting a dive type for a set of environmental conditions is multinomial problem. Unfortunately there are no analytical techniques currently available that take into account serial autocorrelation in the residuals, repeated measures, or non-linearity in the relationships between the response and covariates for multinomial data. The lack of flexibility of the currently available fitting method for multinomial data is illustrated by the poor fit of a multinomial model to the dive type data. The smooth functions for the GAMs, fitted in the `mgcv` library in R (Wood 2006), suggest that responses are in fact non-linear, however the high number of internal knots estimated by the `gam` function produces smooth functions that do not lend themselves to biological interpretation. This is a result of the size of the dataset used in this analysis, whereby results are statistically but not biologically significant, which can be seen in plots of the smooth functions on the scale of the response for the GAMs.

A potentially important feature of the data that cannot be addressed in the framework of the fixed effects models used here, is the correlation between dives by the same individual. The correlograms used to investigate the correlation structure in the residuals of the GAMs suggest low correlation between successive residuals in all dive types. This is surprising since we might expect certain behaviours, such as foraging, to be grouped both in space and time. Correlation appears to be higher in the most common dive type, U dives, which suggests that the correlation structure in the residuals for binary data might be somehow affected by large numbers of zeros in the observed values; but the correlation is still not high enough to be substantially affecting the model fit. A way of finding out whether the correlation plays an important role in model fit would be to incorporate it into the model and see if the fit improves. This can be done either in the context of a mixed model, or in the context of a General Estimating Equation (GEE). The advantage of GEEs is that they can be used to estimate a correlation structure within a specified group of residuals, for example 100 dives, rather than fit a predefined error structure (Hardin & Hilbe 2003). The between-group correlation is assumed to be zero.

4.3. Unravelling the Demise of a Regression Analysis

Reasons why the regression analysis approach failed to explain a high proportion of the variability in the data might be (1) that the dive classification masked any functional dive categories, (2) the classification is meaningful but the RF algorithm performed poorly at assigning dives to the correct classes, (3) the covariates that are most important in explaining patterns in the data are not included in this analysis, or (4) that one or more of the assumptions of the methods used here have been violated, for example independence of dive types.

4.3.1. Dive Classification: Good Idea – Bad Idea

A growing number of studies employ dive classification techniques to aid behavioural analysis of dive profiles, which dates back to Kooyman (1968) (Tinker *et al.* 2007, Fuiman *et al.* 2007, Davis *et al.* 2003, Baechler *et al.* 2002, Shreer & Testa 1996, Schreer & Testa 1995). Originally dive classification was based on maximum depth and dive duration (Kooyman 1968), but dive shape (time vs. depth trace) has also been used extensively as the basis for dive classification since (Baechler *et al.* 2002, Hindell *et al.* 1991). The methods used to classify dives based on dive parameters obtained from bio-logging instruments include statistical methods such as cluster analysis (Tinker *et al.* 2007), principal component analysis (Schreer & Testa 1995), discriminant function analysis (Baechler *et al.* 2002, Schreer & Testa 1995), as well as manual classification (Baechler *et al.* 2002, Schreer & Testa 1996). For Weddell seals Schreer & Testa (1995) found cluster analysis more appropriate than principal component analysis classifying dives into three types, whereas Baechler *et al.* (2002) found as many as seven in harbour seals. Schreer & Testa (1995) also emphasize the need for more efficient and objective classification methods, however irrespective of the classification method, it is imperative that the resulting dive classes both encompass and represent functionally distinct and ecologically meaningful behaviours. The number of classes identified in each analysis appears to depend on the method

employed to classify them (Schreer & Testa 1996), with manual classification resulting in more classes than statistical methods; a finding which does not place much confidence in the classes derived from purely statistical or purely manual classification.

The dive types that are identified vary between studies, classification methods and species, as well as in the corresponding functions attached to each type, but some reoccurring patterns have also emerged. Most large diving vertebrates such as pinnipeds and penguins perform V-shaped dives (V), square-bottom dives with wiggles (W), U-shaped dives, and many also perform square-bottom dives, here termed SQ dives, although no wide ranging assumptions are made regarding the functionality of each dive shape (Baechler *et al.* 2002, Tremblay & Cherel 2000, Hindell *et al.* 1991, Schreer & Testa 1995, 1996). Additional classes used here, R and DR dives, have followed Hindell *et al.* (1991). One important distinction between other studies of dive class and this one is that here U and V dives are considered primarily traveling dives, in contrast to Baechler *et al.* (2002) who mostly associated U dives with foraging dives in harbour seals (*Phoca vitulina*). It is likely that U dives might be traveling dives in southern elephant seals, which forage in a patchy pelagic environment, as they are the most commonly performed dives, however opportunistic feeding should be considered likely during all dive types, irrespective of inferred function. Although animal-borne cameras have been used for ice inhabiting marine vertebrates such as Weddell seals and penguins, direct observations of feeding events have not been coherently related to distinctive dive parameters, nor have observations been made for wide-ranging, primarily pelagic foragers such as elephant seals. Hence, what unites the functionality of dive types discussed in all the studies mentioned here, is that it is speculative.

There is only one study to date in which the use of archival or satellite telemetry data in measuring differences in foraging behaviour has been validated against direct observations of successful feeding events. Tinker *et al.* (2007) used the California sea otter (*Enhydra lutris nereis*) population to show that it is possible to accurately match observations with dive parameters derived from TDR data, and thus predict prey specialization in individual dives for this species. The sea otter is a unique marine mammal in that it brings its prey to the surface to process and consume it. This gives information on capture success and allows for identification of captured prey items as well as prey handling techniques. Six parameters were found to be important in predicting prey specialization in sea otters, which were used to assign diet type to individuals with high accuracy (Tinker *et al.* 2007). Although Tinker *et al.* (2007) focus on the utility of dive data validation in obtaining an index of individual prey specialization; their discussion can be extended to incorporate the use of dive parameters obtained from bio-logging instruments to elucidate different ecological dive functions. It is important to clarify that the success of this validation process is largely due to the one-to-one coupling between behaviour and prey type in sea otters. An analogous comparison would prove much more challenging for species in which a single behaviour has multiple ecological functions. Tinker et al. (2007) describe this scenario as trophic polymorphism. Evidence for a scenario analogous to trophic polymorphism is illustrated here for elephant seals using SQ dives as an example. This might suggest that the current classification of dive shape in southern elephant seals is not a suitable surrogate for ecological functions.

4.3.2. Missing Covariates and Violated Assumptions

Despite the fact that the models fitted in the current analysis have low predictive power, the usefulness of the explanatory variables is not negated, as results suggest that dive type is related to the environmental variables examined here, albeit weakly. Nonetheless is seems likely that important covariates might have been missed. A point of concern when using derived oceanographic variables as covariates in studies like these is that they will almost certainly offer, at least in part, overlapping information. In this study temperature and salinity were used to derive four variables, which were all used in the analysis. Short of only using temperature and salinity as oceanographic variables, there is no good solution to this problem. It has been shown in previous studies that there is a strong diurnal component to maximum dive depth in southern elephant seals (Bennet *et al.* 2001) and that sex plays a potentially important role in determining foraging strategies and hence movement and behaviour, as was shown for southern elephant seals instrumented at Peninsula Valdes (Campagna *et al.* 1999), as well as northern elephant seals (Le Boeuf *et al.* 2000). Even though time of day is not included in this analysis as a covariate, maximum dive depth will have provided some information on time of day given that dive depth and time of day are correlated. We recommend that future analyses include sex, season and time of day explicitly as covariates. Correlograms showed that the correlation between model residuals for dives of the same type is also weak; however if there is a strong correlation between dives of different types then violation of assumptions for the regression approach employed here is a good basis for moving towards an approach that incorporates that correlation such as Markovian type analyses.

4.3.3. Dive Type Uncertainty

In the current analysis uncertainty in the assignment of dives to classes has not been considered. An improvement here would be to incorporate this uncertainty into the modelling process. A way of incorporating the uncertainty in the assignment of dives to classes is to generate many datasets that represent the different realisations of the RF probabilities of a dive belonging to each type. This can be done as follows: for each dive in the dataset, a class is sampled at random from the RF probabilities for that dive, thus generating a data point. This is done for the whole dataset and repeated, for example 1000 times, thus generating 1000 datasets. The same analysis can then be performed for each dataset. A disadvantage here is that this is a computationally intensive process.

4.4. Alternative Approaches to Dive Type Analysis

Having reviewed the methods available for dealing with these challenges of behavioural data obtained from SRDLs in a regression framework, below are some ideas on how to tackle the problem in an integrated fashion beyond the realm of regression.

Thus far we have employed a method that assumes dive types to be independent. If dive types are in fact related, and two or more dive types are likely to be associated, then the results obtained from the regression analysis are likely to be flawed. A first option for exploring whether there is a relationship between dive types is to extrapolate from the existing hypotheses that are speculated to be true from past work. For example, U and V dives are assumed to be travelling dives therefore the displacement during these dive types should be greater than for non-travelling dives, or that W and DR dives are related because animals rest after active foraging, and also improve in body condition. Challenges involved with the displacement example come from having to calculate it based on Argos locations, as these are not exact but interpolated. Improvement in physical condition can be inferred used weight gain derived from drift rate analyses, however weight gain is not strictly coupled with a positive change in physical condition in the animal. An increase in the fat content of the animal will make it more buoyant whereas an increase in lean tissue will make it less buoyant, even though both represent an improvement in body condition. There are differences between the sexes in how the body composition of animals changes as they gain weight (Beck *et al.* 2003). It is more important for females to put on fat as it is a readily available energy store for feeding young, whereas males need to grow a lot more so it is to their advantage to put on lean tissue (Beck *et al.* 2003).

A second option is to quantify the relationship between dive dives as transition probabilities and incorporate them into a Markovian type analysis, whereby the aim is to model the relationship between transition probabilities and the covariates. In this case, the type of dive performed by an animal is regarded as a discrete-time stochastic process, whose distribution is conditional on the history of the process and each observation is determined by the observation preceding it. Hidden Markov Models (HMM) extend this type of analysis by allowing for measurement errors in observing the response. HMMs assume two stochastic processes: the underlying unobserved process and the observed process. The observed process is regarded as a distortion of the underlying process due to noise, and inferences regarding the underlying process are drawn from the observed process (Cappé *et al.* 2005). This type of approach is likely well suited to predicting dive type as a response to environmental conditions. HMMs might be especially useful if dive classes are assumed to be surrogates for the underlying unobserved pattern in diving behaviour. An assumption of HMMs is that the distribution of time spent in each dive type belongs to the exponential family. If this is not the case for dive type data then a semi-Markov model offers the possibility of less restrictive assumptions (Cappé *et al.* 2005).

4.5. Maximum Dive Depth as a Response to Bathymetry in SQ dives

McConnell *et al.* (1992) first suggested for female elephant seals instrumented at South Georgia, that dives occurring in proximity to the continental shelf are benthic dives during which feeding occurs at an increased rate. This was inferred from reduced duration of these dives, which was interpreted as a reduction in aerobic dive limit (ADL) due to increased foraging activity. Benthic diving behaviour has also been described in rockhopper penguins in the Kerguelen Archipelago, during which birds perform shorter, more efficient and more active “square-wave” dives that closely follow the bathymetry (Tremblay & Cherel 2000). Bailleul *et al.* (2007) have recently shown that for southern elephant seals instrumented at Kerguelen Islands, benthic dives on the

Antarctic continental shelf are associated with increased foraging activity and improved body condition, based on track sinuosity and dive density. What has been shown in the context of the current study is that although there is substantial individual variation in the occurrence of SQ dives, at least some animals dive to the seabed. For those animals there is a statistically significant positive relationship between maximum dive depth and bathymetry. It is worth mentioning that the two animals displaying this behaviour in the current analysis were males. It has been shown that males have greater diving capabilities than females, and that this is reflected in sex differences in foraging behaviour (Campagna *et al.* 1999). This suggests that benthic foraging might be a behaviour more commonly employed by males, and that SQ dives should not be exclusively related to benthic foraging.

Although most of the SQ dives occurred over continental shelf waters, a proportion of them did not. If we attach a benthic feeding functionality to SQ dives in shelf waters, this suggests either that SQ dives occurring over deep water have been misclassified, or that SQ dives are multifunctional, in which case the class itself does not correspond to a distinct function and will not be useful in predicting distinct behaviours. An analysis of the time of day of SQ dives performed in pelagic waters might shed light on their functionality. Bennet *et al.* (2001) have shown that maximum depth shows a strong diurnal variation, with nighttime dives being shallower than daytime dives, corresponding to the vertical diurnal movement of prey within the water column. If these pelagic SQ dives are performed at night, they might represent a different type of foraging, for example ambush dives, in contrast to versus pursuit dives, which might be speculated to correspond to W dives. This hypothesis can be easily tested in further analysis. This example of SQ dives casts uncertainty upon the functional distinction between the dive classes assumed here, as it illustrates an uncoupling of the response and explanatory variables measured.

4.6. Conclusions

Regression analysis of dive type in relation to environmental conditions was not found to be an appropriate method for analysing abstracted behavioural dive data, illustrated by the lack of satisfactory fit of the models to the data. There was a weak relationship between the response and the covariates included in this analysis, however it was not sufficient to enable dive type predictions. Although regression has not been exhausted as a method of capturing the relationship between dive type and environmental variables, it seems unlikely that it will provide further insight. The fact that SQ dives might represent multiple behaviours in southern elephant seals warrants a closer investigation of the usefulness of dive type in identification of individual ecological functions and its subsequent analysis as a response to environmental variables. Despite the substantial analytical challenges, preliminary analyses such as this are likely to lead to improvements in interpretability of abstracted behavioural data, and to the development of a standardised method for widespread application to such abstracted behavioural data, obtained in abundance via satellite telemetry. Gaining an understanding of how the spatial and temporal variability in environmental conditions in the Southern Ocean affect predator diving, may be used to predict how this behaviour could change in response to oceanographic variability, as a result of climatic variation.

5. References

Agresti, A. 1996. An Introduction to Categorical Data Analysis. Wiley Series in Probability and Statistics. John Wiley & Sons, United States of America
ISBN 0-471-11338-7

Argos 2007. www.argos-system.org

Bailleul, F., Charrassin, J.B. & Monestiez, P., Roquet, F., Biuw, M. & Guinet, C. 2007. Successful foraging zones of southern elephant seals from the Kerguelen Islands in relation to oceanographic conditions. *Philosophical Transactions of the Royal Society B.* Published on-line doi:10.1098/rstb.2007.2109

Beck, C.A., Bowen, W.D. & Iverson, S.J. 2003. Sex differences in the seasonal patterns of energy storage and expenditure in a phocid seal. *Journal of Animal Ecology,* 72: 280–291

Bennet, K. A., McConnell, B.J. & Fedak, M.A. 2001. Diurnal and seasonal variations in the duration and depth of the longest dives in southern elephant seals (Mirounga leonina): Possible physiological and behavioural constraints. *The Journal of Experimental Biology.* 204: 649–662

Biuw, M. et al. 2007. Variations in behavior and condition of a Southern Ocean top predator in relation to *in situ* oceanographic conditions. *Proceedings of the National Academy of Sciences* 104(34): 13705–13710

Biuw, M., Patterson, T., Thums, M., Hindell, M. & Fedak, M. Dives through Random Forests: Yet another classification method for 2-dimensional dive profiles. *in prep.*

Biuw, M., McConnell, B., Bradshaw, C.J.A., Burton, H. & Fedak, M. 2003. Blubber and buoyancy: monitoring the body condition of free-ranging seals using simple dive characteristics. *The Journal of Experimental Biology*, 206: 3405-3423

Boehlert, G.W., Costa, D.P., Crocker, D.E., Green, P., O'Brien, T., Levitus, S. & Le Boeuf, B.J. 2001. Autonomous pinniped environmental samplers: Using instrumented animals as oceanographic data collectors. *Journal of Atmospheric and Oceanographic Technology* 18: 1882-1893

Boehme, L, Thorpe, S.E., Meredith, M.P., Biuw, M. & Fedak, M. 2007. Mothly snapshots of the ACC frontal system in the South Atlantic by merging Argo and animal-borne sensors. *Journal of Marine Science, In press*

Le Boeuf, B.J., Crocker, D.E., Costa, D.P., Blackwell, S.B., Webb, P.M. & Houser, D.S. 2000a. Foraging ecology in northern elephant seals. *Ecological Monographs,* 70(3): 353-382

Le Boeuf, B.J., Crocker, D.E., Grayson, J., Gedamke J., WEBB, P.M., Blackwell, S.B. & Costa, D.P. 2000b. Respiration and heart rate at the surface between dives in southern elephant seals. *The Journal of Experimental Biology*, 203: 3265–3274

Breiman, L. 2001. Random Forests. *Machine Learning*, 45: 5–32

Campagna, C., Fedak, M. & McConnell, B. 1999. Post-breeding distribution and diving behaviour of adult male southern elephant seals from Patagonia. *Journal of Mammalogy* 80(4):1341-1352

Cappé, O., Moulines, E. & Rydén, T. 2005. Inference in Hidden Markov Models. Springer Series in Statistics. Springer, United States of America.
ISBN 0-387-40264-0

Charrassin, J.B., Park, Y.H., Le Maho, Y. & Bost, C.A. 2002. Penguins as oceanographers unravel hidden mechanisms of marine productivity. *Ecology Letters* 5: 317–319

Crocker, D.E., Costa, D.P., Le Boeuf, B.J., Webb, P.M. & Houser, D.S. 2006. Impact of El Niño on the foraging behavior of female northern elephant seals. *Marine Ecology Progress Series* 309: 1–10

Davis, R.W., Fuiman, L.A, Williams, T.M., Horning, M. & Hargey, W. 2003. Classification of Weddell seal dives based on 3-dimensional movements and video-recorded observations. *Marine Ecology Progress Series*, 264: 109–122

ETOPO5 Grid-cell registration. 1988. U.S. Department of Commerce. National Oceanic and Atmospheric Administration. National Environmental Satellite, Data, and Information Service. National Geophysical Data Center.
http://www.ngdc.noaa.gov/mgg/global/etopo5.HTML

ETOPO2v2g Grid-cell registration. 2007. U.S. Department of Commerce. National Oceanic and Atmospheric Administration. National Environmental Satellite, Data, and Information Service. National Geophysical Data Center.
http://www.ngdc.noaa.gov/mgg/fliers/06mgg01.html

Faraway, J.J. 2006. Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models. CRC Press. ISBN 1-58488-424-X

Fedak, M., Lovell, P. & Grant, S.M. 2001. Two approaches to compressing and interpreting time-depth information as collected by time-depth recorders and satellite-linked data recorders. *Marine Mammal Science*, 17(1): 94-110

Fedak, M., Lovell, P., McConnell, B. & Hunter, C. 2002. Overcoming the Constraints of Long Range Radio Telemetry from Animals: Getting More Useful Data from Smaller Packages. *Integrative and Comparative Biology* 42(1): 3–10

Fedak, M.A. 2003. Marine animals as platforms for oceanographic sampling: a "win/win" situation for biology and operational oceanography. *Memoirs of National Institute of Polar Research*, Special Issue 58: 133–147, 2004

Fuiman, L.A., Maddena, K.M., Williams, T.M., Davis, R.W. 2007. Structure of foraging dives by Weddell seals at an offshore isolated hole in the Antarctic fast-ice environment. *Deep-Sea Research II*, 54: 270–289

Hardin, J.W. & Hilbe, J.M. 2002. Generalized Estimating Equations. Chapman & Hall/CRC. ISBN 1-58488-307-3

Heywood, K.J. & King, B.A. 2002.Water masses and baroclinic transports in the South Atlantic and Southern Oceans. *Journal of Marine Research* 60(5): 639-367

Hindell, M.A., Bradshaw, C.J.A, Sumner, M.D., Michael, K.J. & Burton, H.R. 2003 Dispersal of female southern elephant seals and their prey consumption during the austral summer: relevance to management and oceanographic zones. *Journal of Applied Ecology* 40: 703–715

Hindell, M.A., Slip, D.J. & Burton, H.R. 1991. The diving behaviour of adult male and female southern elephant seals, *Mirounga leonina* (Pinnipedia, Phocidae) *Australian Journal of Zoology*, 39(5): 595-619

Jonker, F.C. & Bester, M. N. 1998. Seasonal movements and foraging areas of adult southern female elephant seals, *Mirounga leonina*, from Marion Island. *Antarctic Science* 10: 21-30

Kara, A.B., Rochford, P.A., Hurlburt, H.E. 2000. An optimum definition for ocean mixed layer depth. *Journal of Geophysical Research*, 105: 16803-16821

Kooyman, G. L. 1968. An analysis of some behavioral and physiological characteristics related to diving in the Weddell seal. Pages 227-261 *in* G. A. Llano and W. L. Schmitt, eds. Biology of the Antarctic seas. Antarctic Research Series. Volume 3. American Geophysical Union, Washington, DC.

Lesage, V., Hammill, M.O., and Kovacs, K.M. 1999. Functional classification of harbor seal (*Phoca vitulina*) dives using depth profiles, swimming velocity, and an index of foraging success. *Canadian Journal of Zoology*, 77: 74–87.

Liaw, A. and Wiener, M. (2002). Classification and Regression by `randomForest`. R News 2(3): 18-22.

Lydersen, C., Nost, O.A., Lovell, P., McConnell, B.J., Gammelsrod, T., Hunter, C., Fedak, M. & Kovacs, K.M. 2002. Salinity and temperature structure of a freezing Arctic fjord – monitored by white whales (*Delphinapterus leucas*). *Geophysical Research Letters* 29(23): 2119-2123

McConnell, B.J, Chambers, C & Fedak, M.A.1992. Foraging ecology of southern elephant seals in relation to the bathymetry and productivity of the Southern Ocean. *Antarctic Science* 4(4): 393-398

McMahon, C.R., Bradshaw, C.J.A. & Hays, G.C. 2007. Satellite tracking reveals unusual diving characteristics for a marine reptile, the olive ridley turtle *Lepidochelys olivacea. Marine Ecology Progress Series* 329: 239–252

Murrell, P. 2006. R Graphics. CRC Press USA. ISBN 1-58488-486-X

Myers, A.E., Lovell, P. & Hays, G. 2006. Tools for studying animal behaviour: validation of dive profiles relayed via the Argos satellite system. *Animal Behaviour*, 71: 989–993

Pinheiro, J., Bates, D., DebRoy, S. & Sarkar D. the R Core team. (2007). nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-83.

R Development Core Team (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
ISBN 3-900051-07-0, URL http://www.R-project.org.

Schreer, J.F., O'Hara Hines, R.J. & Kovacs, K.M. 1998. Classification of dive profiles: A comparison of statistical clustering techniques and unsupervised artificial neural networks. *Journal of Agricultural, Biological, and Environmental Statistics*, 3(4): 383-404

Schreer, J.F. & Testa, J.W. 1996. Classification of Weddell seal diving behaviour. *Marine Mammal Science*, 12(2): 227-250

Schreer, J.F. & Testa, J.W. 1995. Statistical classification of diving behaviour. *Marine Mammal Science*, 11(1): 85-93

Tinker, M.T., Costa, D.P, Estes, J.A. & Wieringa, N. 2007. Individual dietary specialization and dive behaviour in the California sea otter: Using archival time–depth data to detect alternative foraging strategies. *Deep-Sea Research II*, 54: 330–342

Tremblay, Y. & Cherel, Y. 2000. Benthic and pelagic dives: a new foraging behaviour in rockhopper penguins. *Marine Ecology Progress Series* 204: 257-267

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, NewYork. ISBN 0-387-95457-0

Wall, S.M., Bradshaw, C.J.A., Southwell, C.J., Gales, N.J. & Hindell, M.A. 2007. Crabeater seal diving behaviour in eastern Antarctica. *Marine Ecology Progress Series* 337: 265–277

Wilson, R.P., and Wilson, M.P. 1995. The foraging behaviour of the African penguin. *In* The penguins ecology and management. *Edited by* P. Dann, I. Norman, and P. Reilly. Surrey Beatty and Sons, Sidney. pp. 244–265.

Wood, S. 2006. Generalized Additive Models: An Introduction with R. CRC Press USA. ISBN 1-58488-474-6

6. Appendix A - Methods

Table 1. Deployment and track details for individuals included in the South Georgia dataset.

Individual Reference	Year of Deployment	Sex	Track length after filtering (days)
ct8-Rudolph-05	2005	M	105
ct8-Jason-05	2005	M	31
ct1-Donner-04	2004	F	161
ct1-Dancer-04	2004	F	17
ct8-2849-05	2005	F	35
ct8-Bernt-05	2005	M	36
ct1-Dasher-04	2004	F	26
ct1-Vixen-04	2004	F	202
ct1-Blitzen-04	2004	F	207
ct8-Undine-05	2005	F	32
ct8-Aspasia-05	2005	F	32
ct8-26626-05	2005	F	26
ct8-Carlita-05	2005	F	31
ct8-Rosita-05	2005	F	33
ct8-28494-05	2005	F	114
ct1-Comet-04	2004	F	36
ct8-Soern-05	2005	M	113

Table 2. Dive parameters used in random forest classification of dives into dive types.

Dive Parameter	Definition
Dive duration	Duration of a dive (sec)
Maximum dive depth	Depth at inflection point with maximum depth
Bottom depth	Bathymetry (m)
pr1.4	Proportion of dive time spent between inflection points 1 and 4
pr1.3	Proportion of dive time spent between inflection points 1 and 3
pr2.4	Proportion of dive time spent between inflection points 2 and 4
signs	Direction of swimming between inflection points 1-4 (e.g. 'DUD' means the seal swam down between point 1-2, up between point 2-3 and again down between point 3-4.)
des	Descent rate
asc	Ascent rate
Slope1	The rate of descent or ascent between point 1-2
Slope2	The rate of descent or ascent between point 2-3
Slope3	The rate of descent or ascent between point 3-4
Mean1.4	Mean rate of vertical change between point 1-4 (i.e. average of the three above)
SD.1.4	Standard deviation of the three above
Mean1.3	Same as above, but excluding point the segment between point 3-4
SD.1.3	Same as above, but excluding point the segment between point 3-4
Mean2.4	Same as above, but excluding point the segment between point 1-2
SD.2.4	Same as above, but excluding point the segment between point 1-2

Table 3. Random forest classification confusion matrix and class error for each type as well as overall out of bag (OOB) error rate

			OOB estimate of error rate 3.6%				
Dive Type	SQ	U	V	W	DR	R	Class Error
SQ	13	6	0	0	0	0	0.3158
U	1	1549	2	0	2	3	0.0051
V	0	8	27	0	1	3	0.3077
W	0	8	0	982	0	15	0.0229
DR	0	12	0	0	114	0	0.0952
R	0	3	0	44	0	207	0.1850

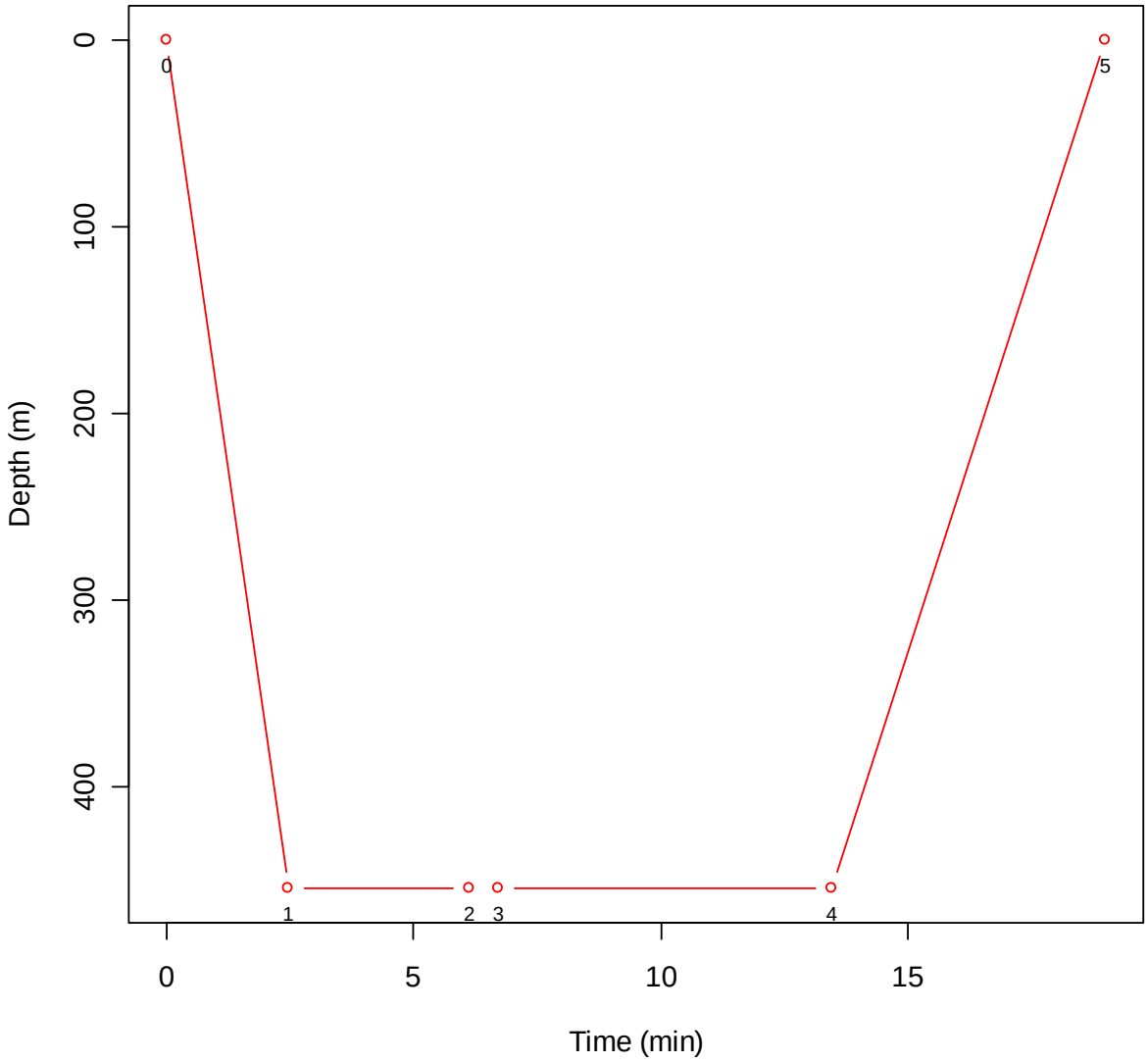

Figure 1. Sample of a square bottom dive classified here as a SQ dive.

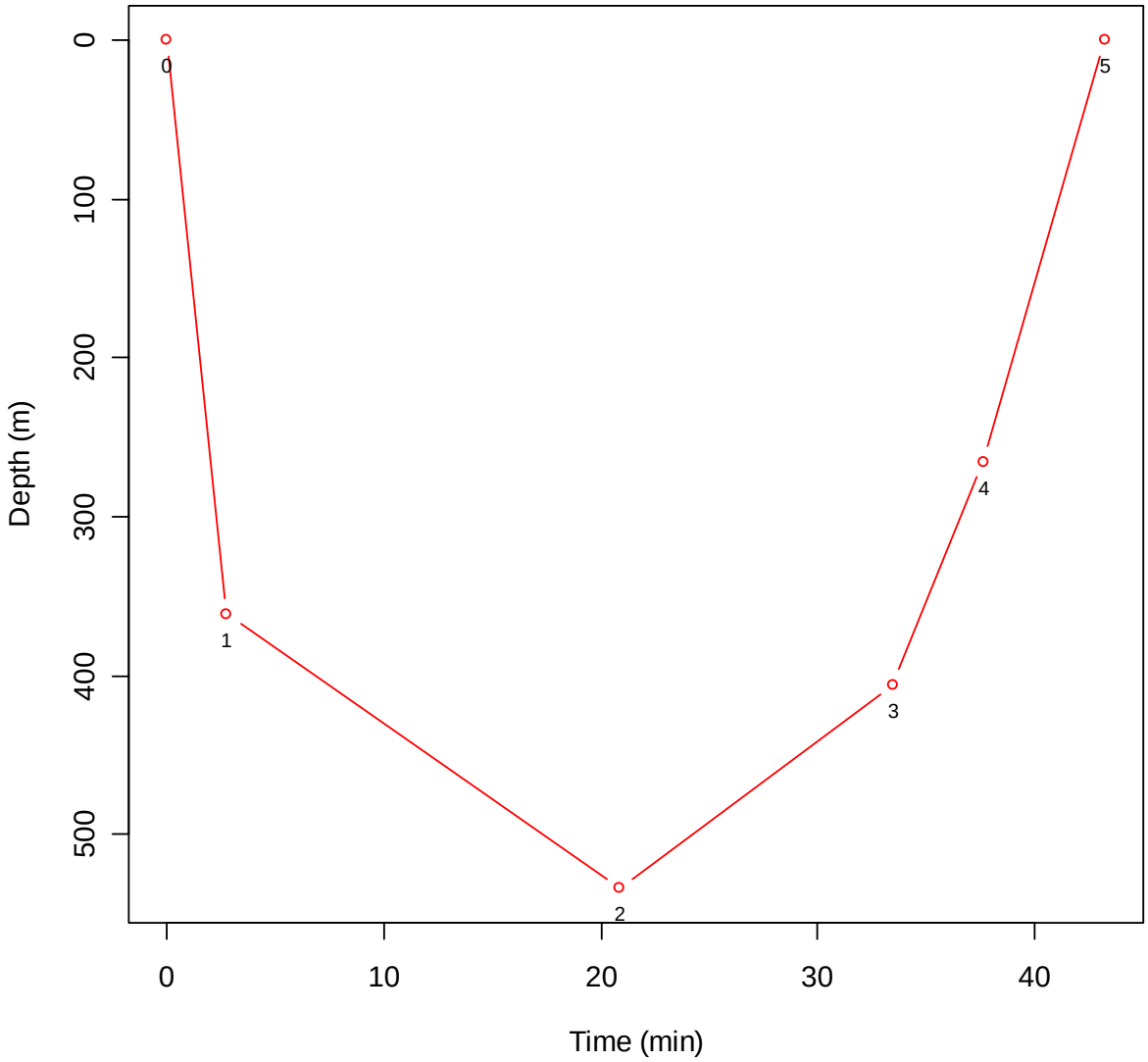

Figure 2. Sample of a U-shaped dive classified here as a U dive.

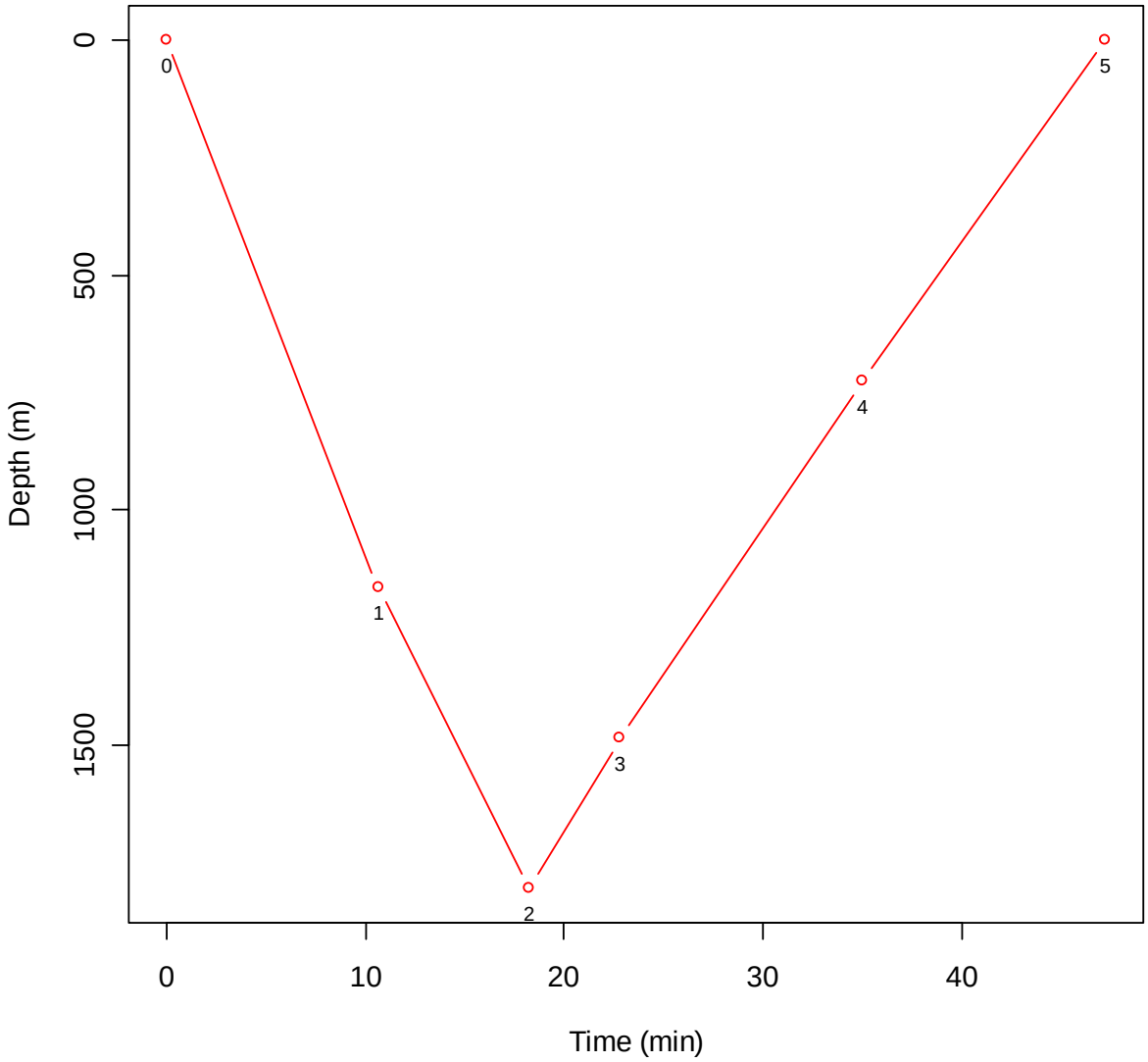

Figure 3. Sample of a V-shaped dive classified here as a V dive.

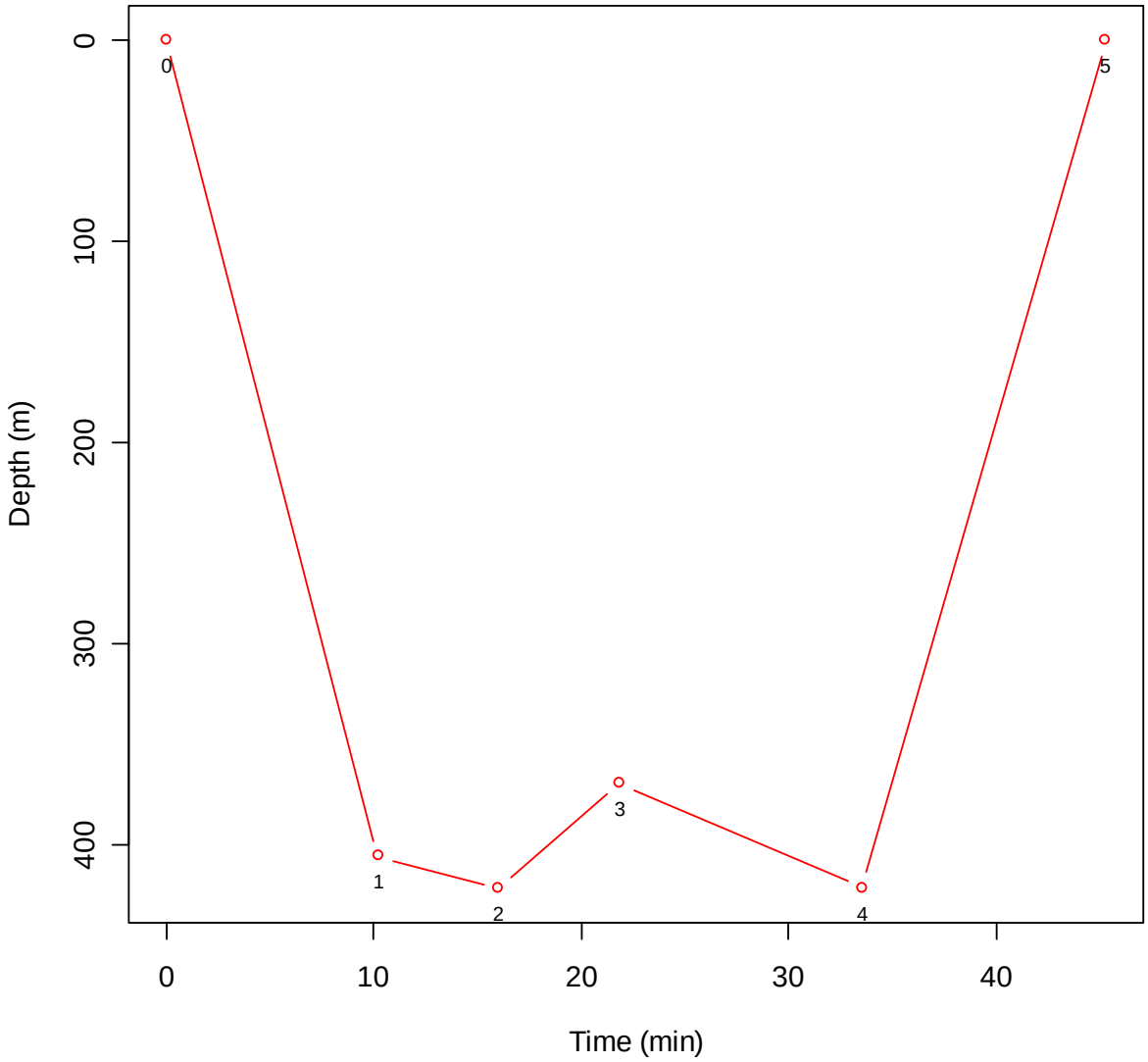

Figure 4. Sample of a square bottom dive with wiggles classified here as a W dive.

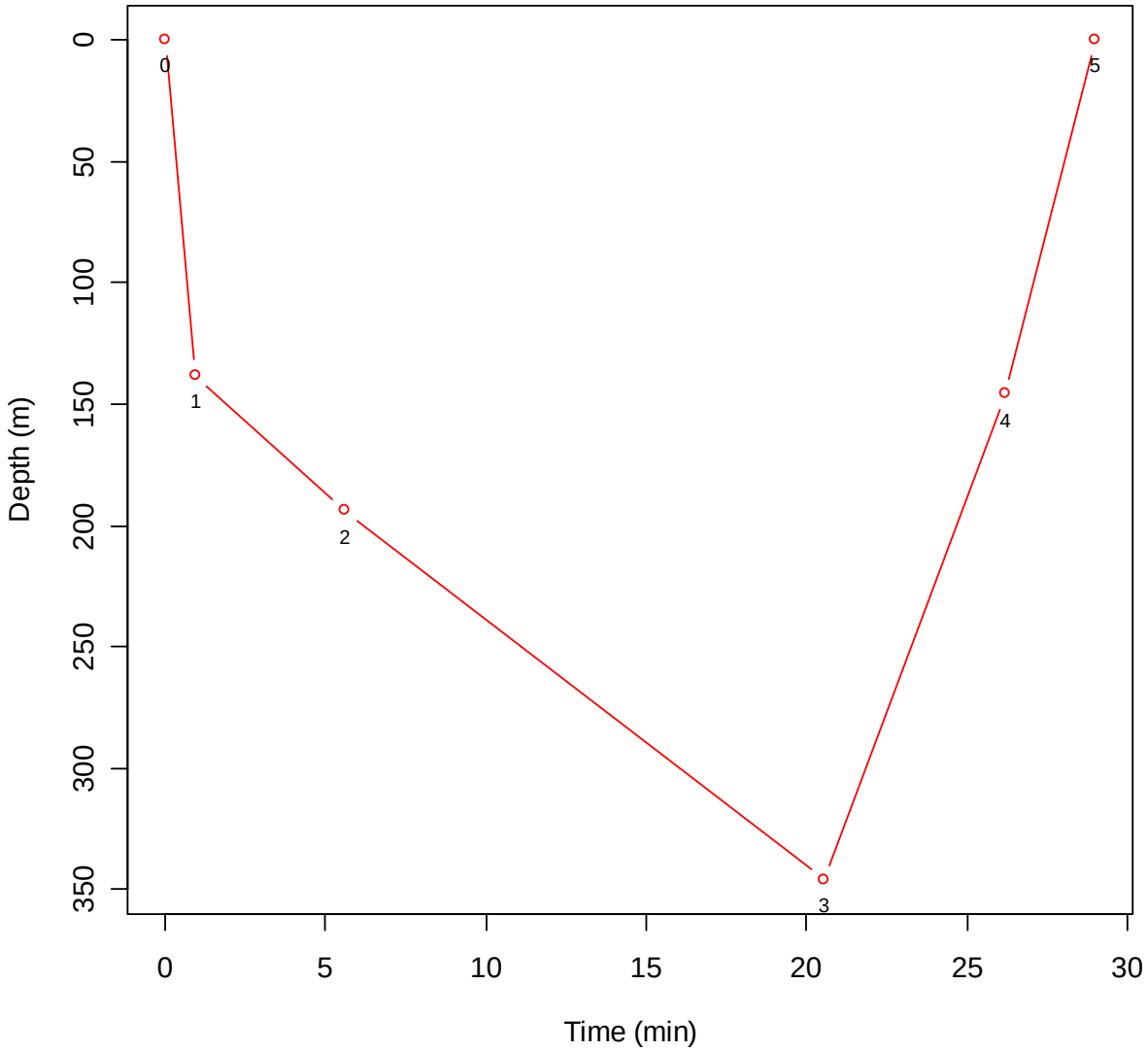

Figure 5. Sample of a drift dive classified here as a DR dive.

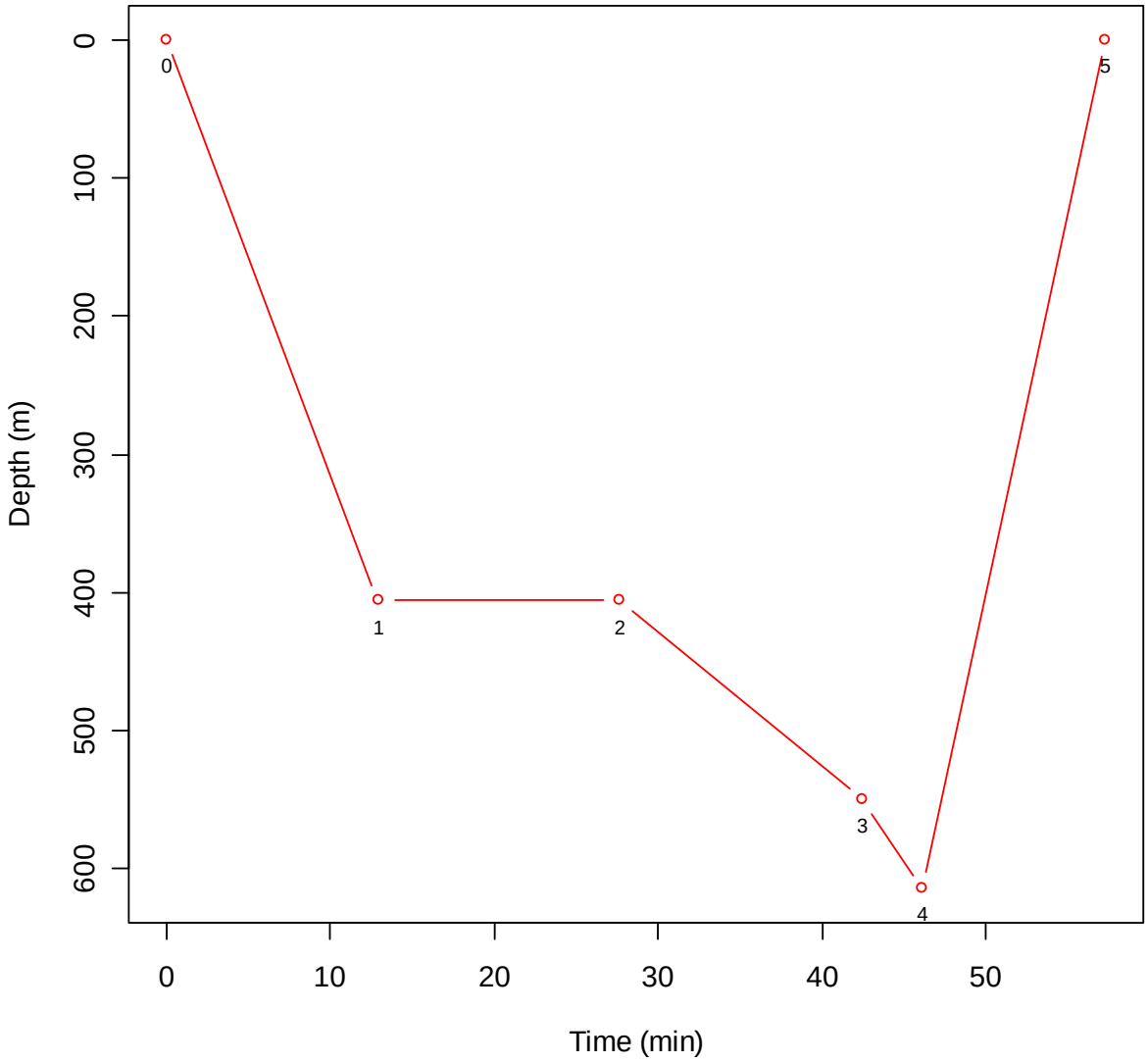

Figure 6. Sample of a root shaped dive classified here as an R dive.

7. Appendix B - Results

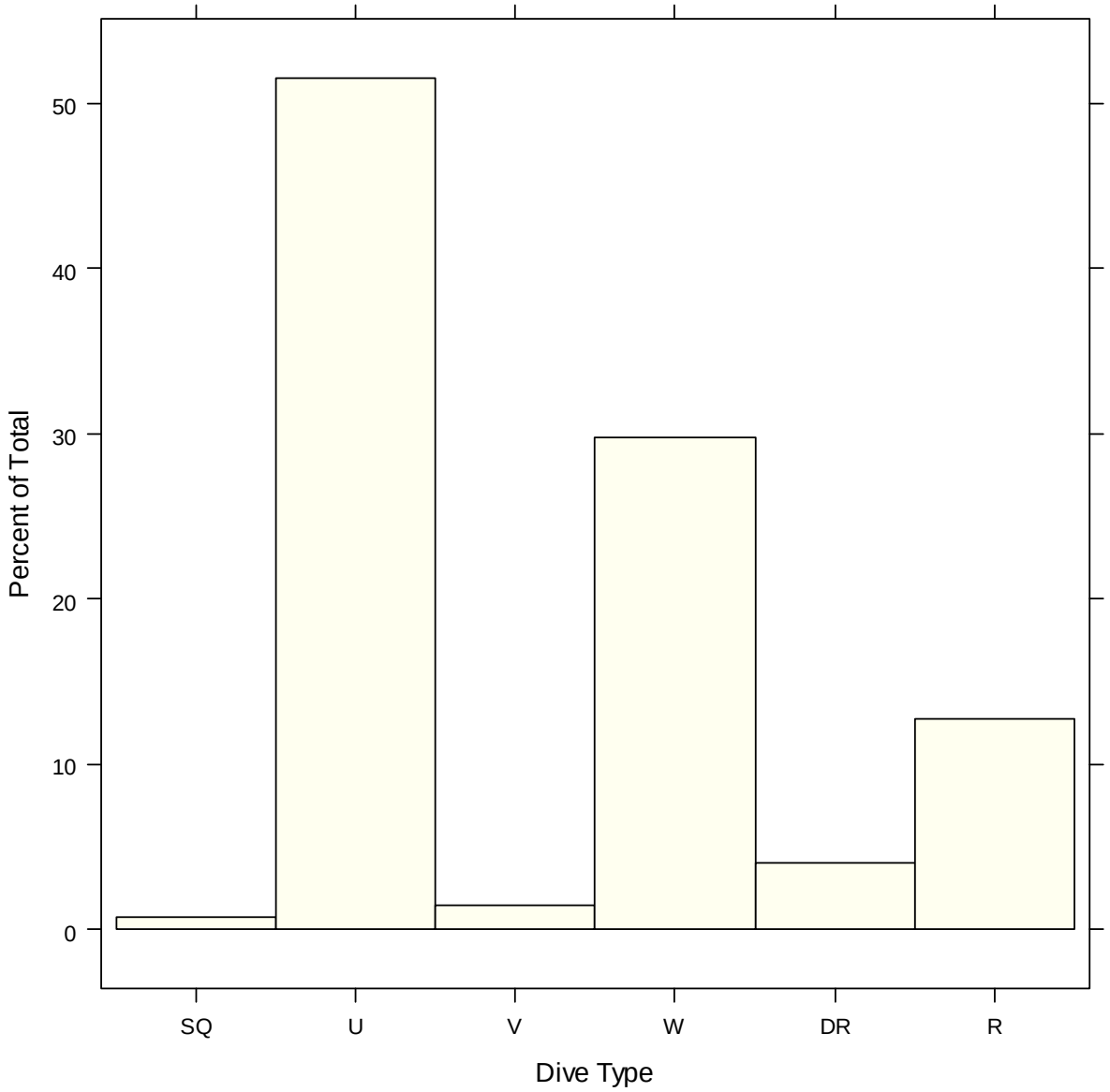

Figure 7. Overall relative occurrence of dive type across individual southern elephant seals from South Georgia in 2004 and 2005

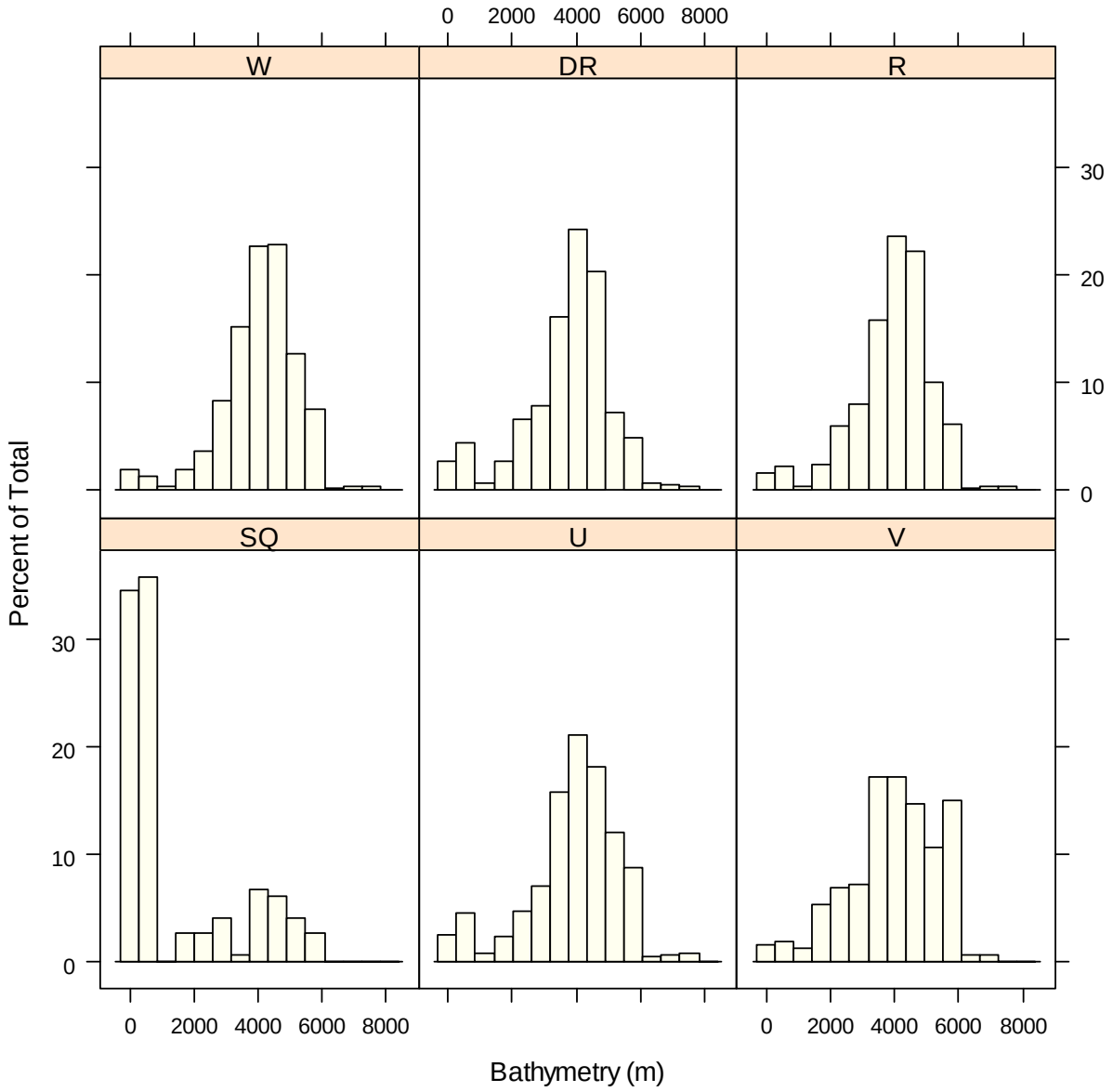

Figure 8. Proportion of total dives per dive type that occured in each bathymetric stratum for each dive type.

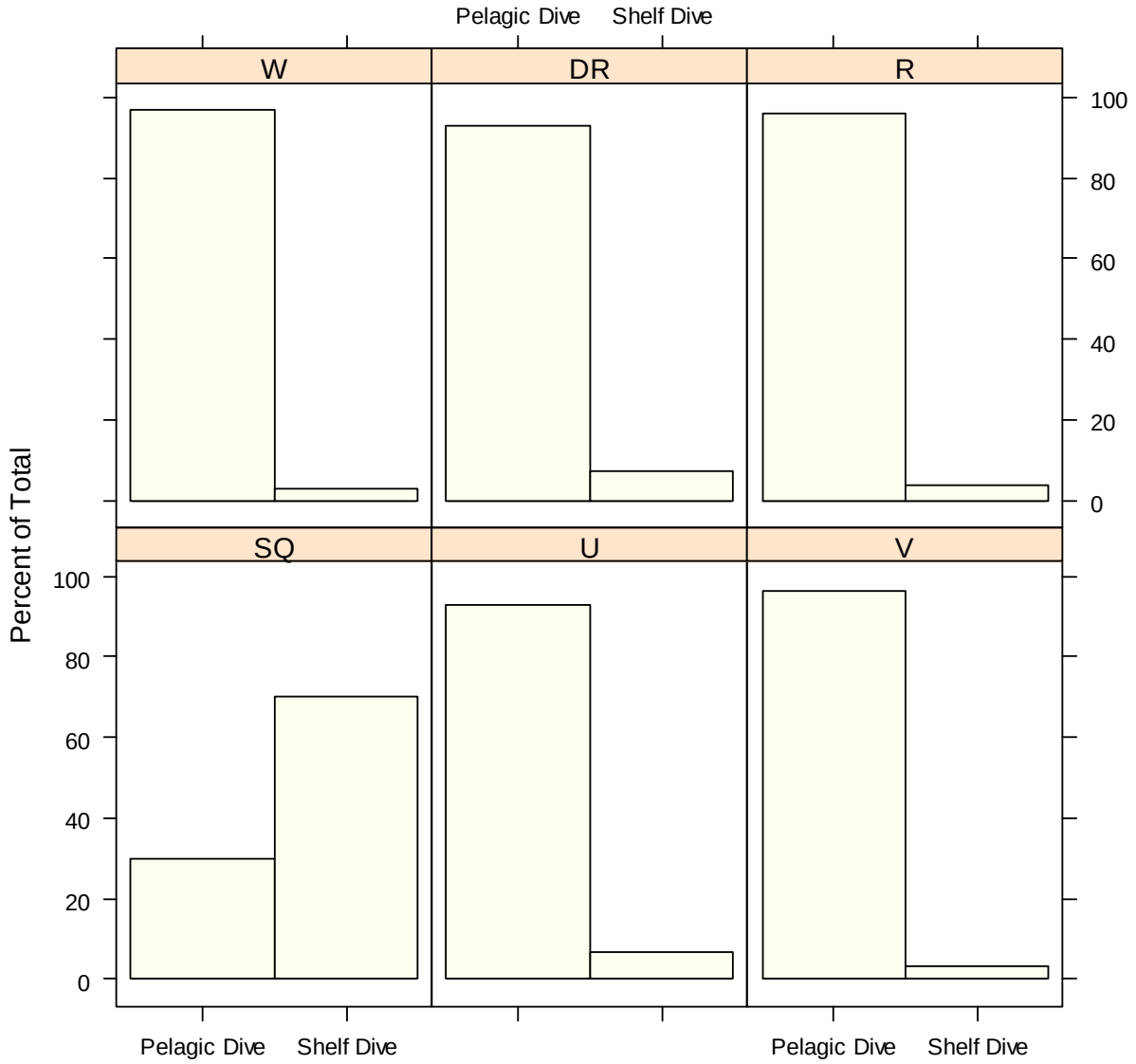

Figure 9. Proportion of total dives per dive type that occurred in continental shelf waters [shelf waters (1), pelagic waters (0)] for each dive type.

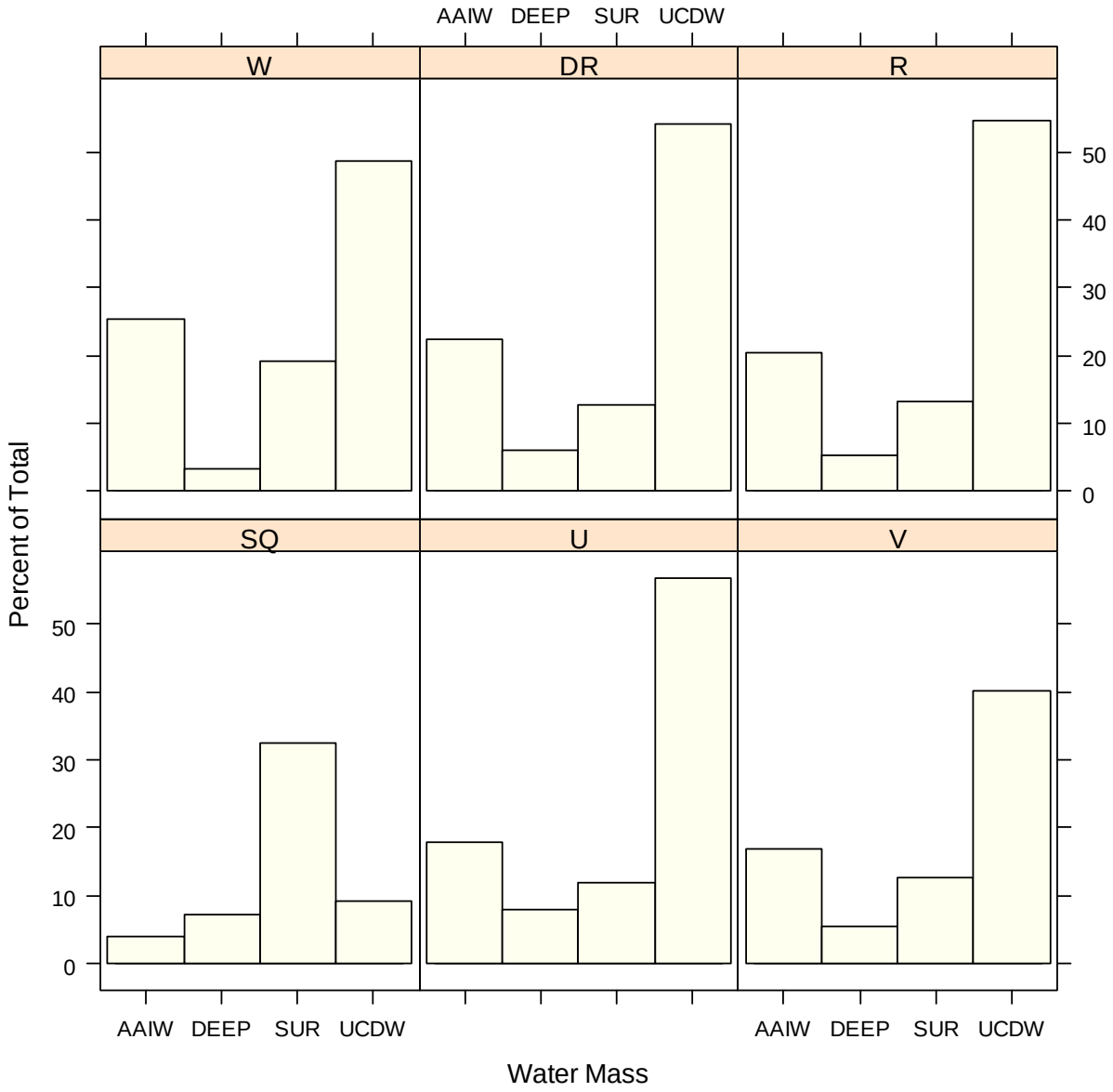

Figure 10. Proportion of total dives per dive type that occurred in each water mass (AAIW: Antarctic Intermediate Water, DEEP: Bottom Water, SUR: Surface Water, UCDW: Upper Circumpolar Deep Water).

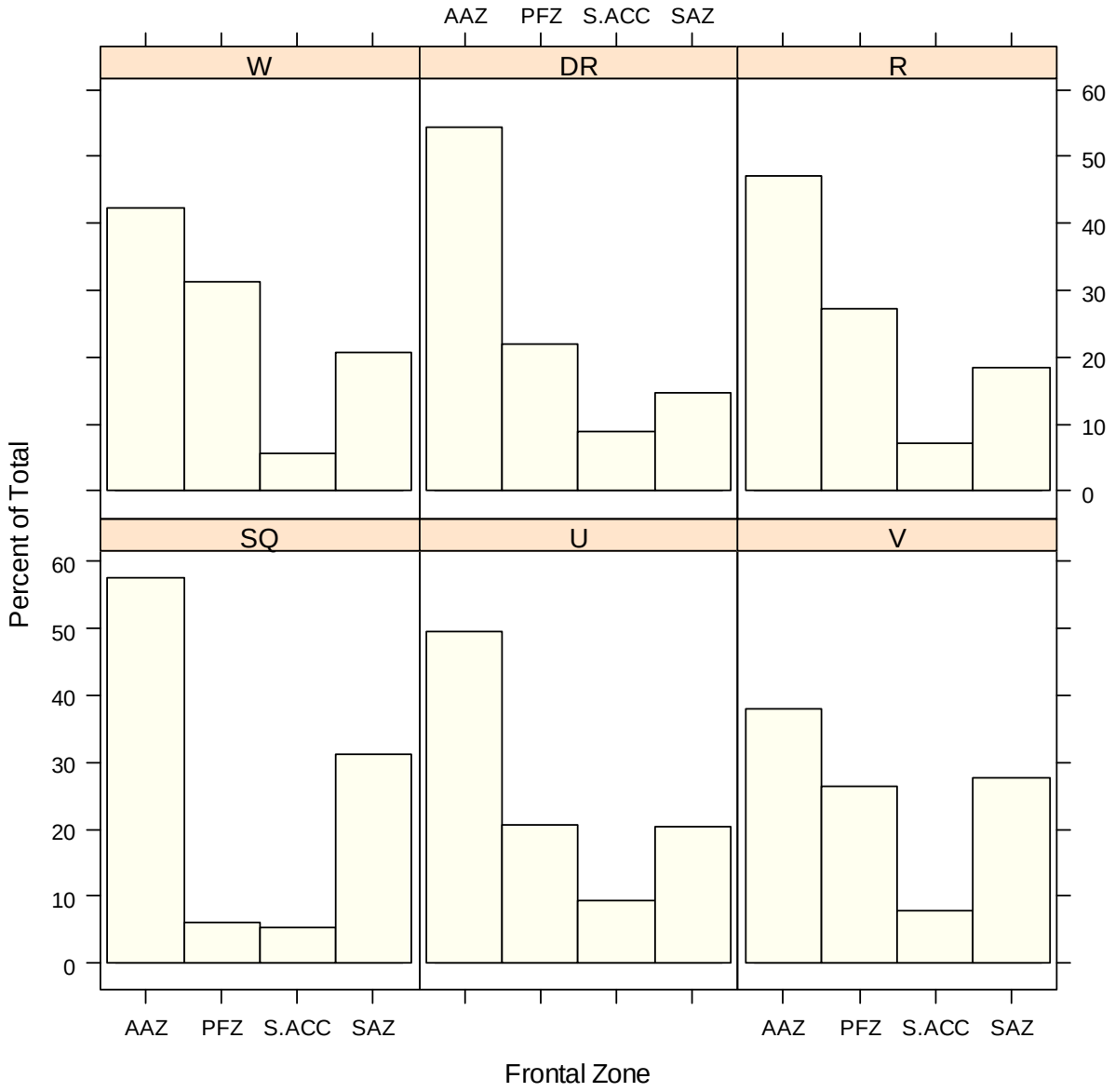

Figure 11. Proportion of total dives per dive type that occurred in each frontal zone (S.ACC: South of Antarctic Circumpolar Current, AAZ: Antarctic Current, PFZ: Polar Frontal Zone, SAZ: Subantarctic Zone).

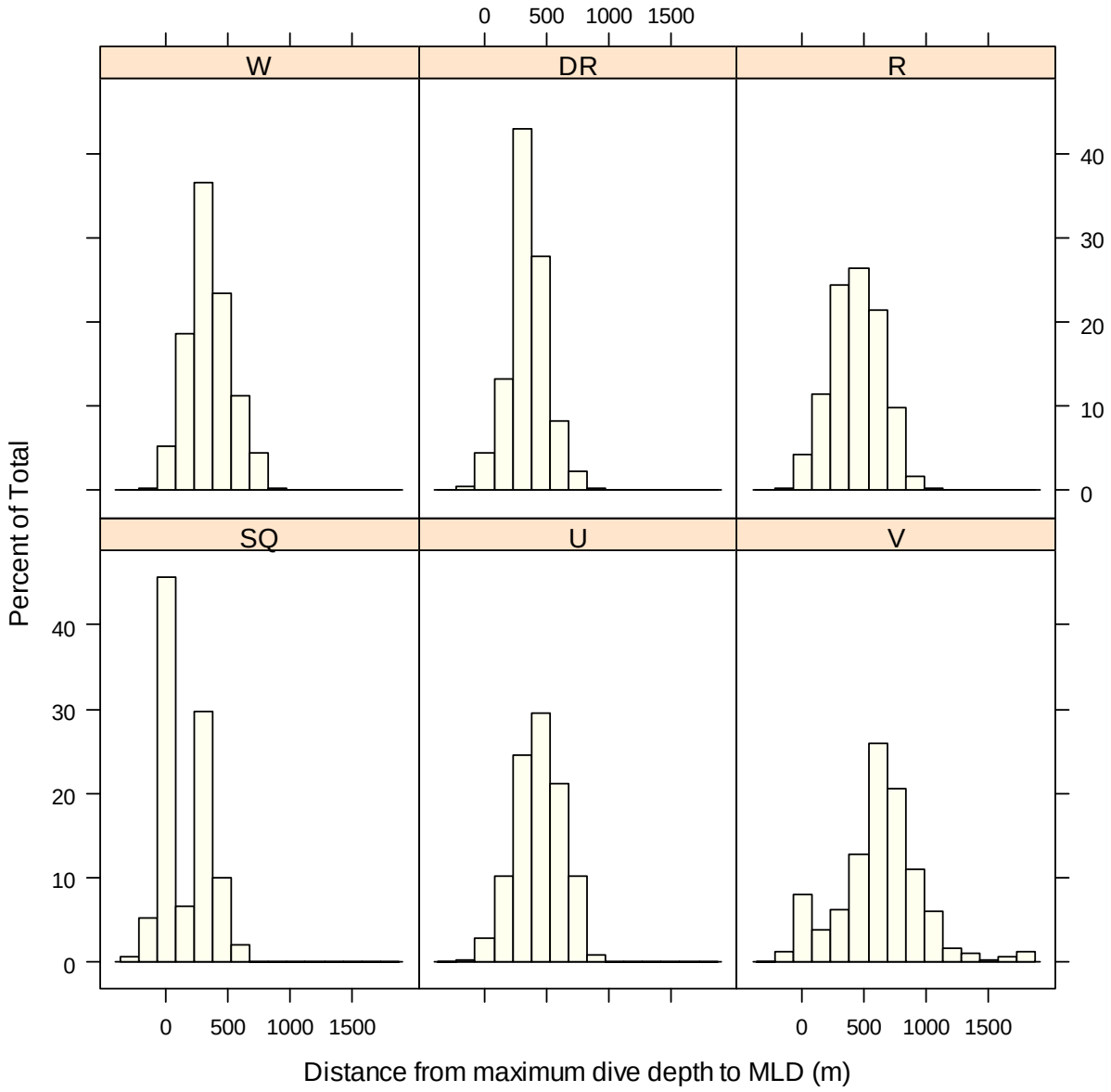

Figure 12. Distribution of distances (m) from maximum dive depth to Mixed Layer Depth (MLD) for each dive type, presented as the proportion of total dives per dive type.

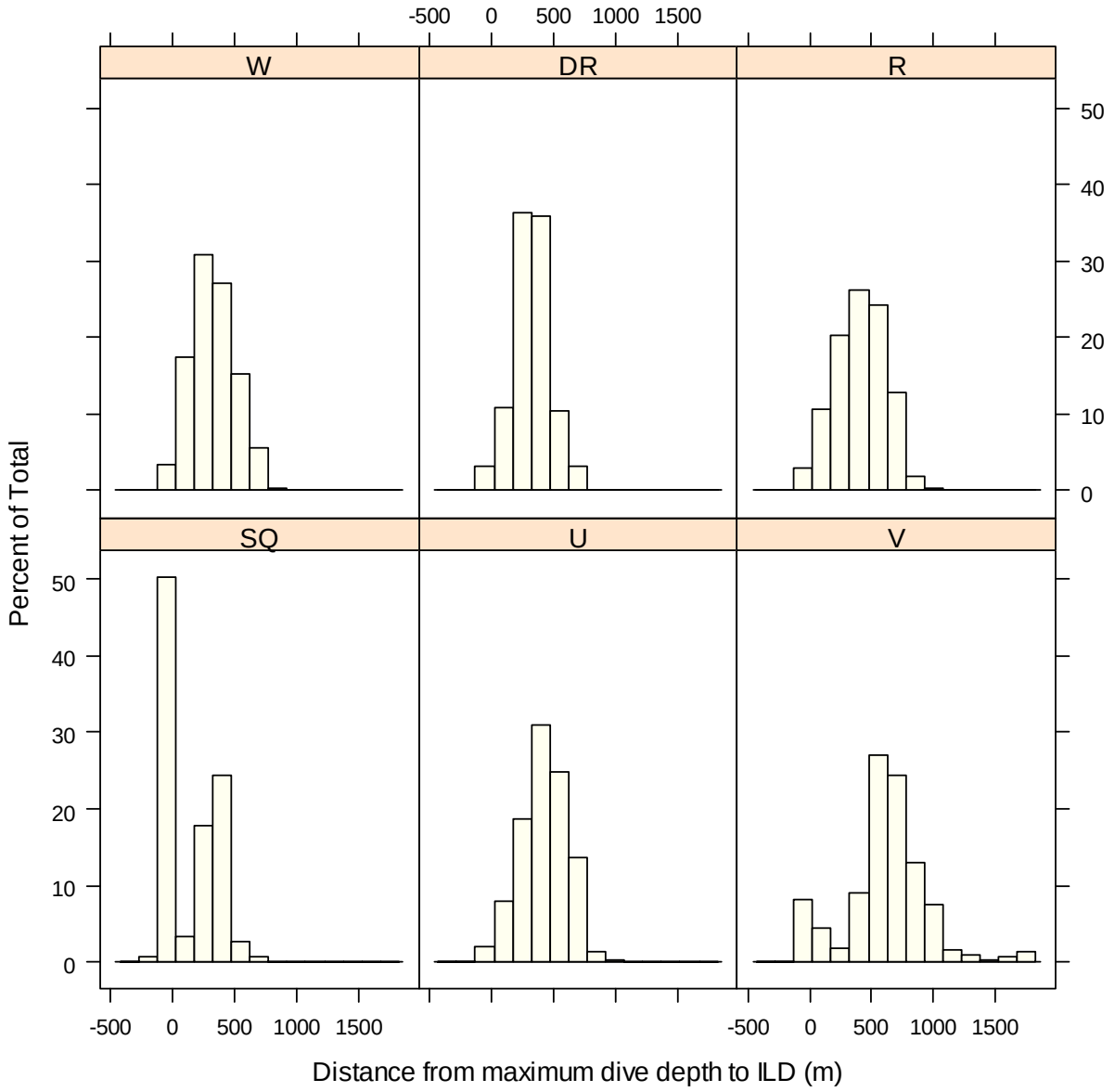

Figure 13. Distribution of distances (m) from maximum dive depth to Mixed Layer Depth (ILD) for each dive type, presented as the proportion of total dives per dive type.

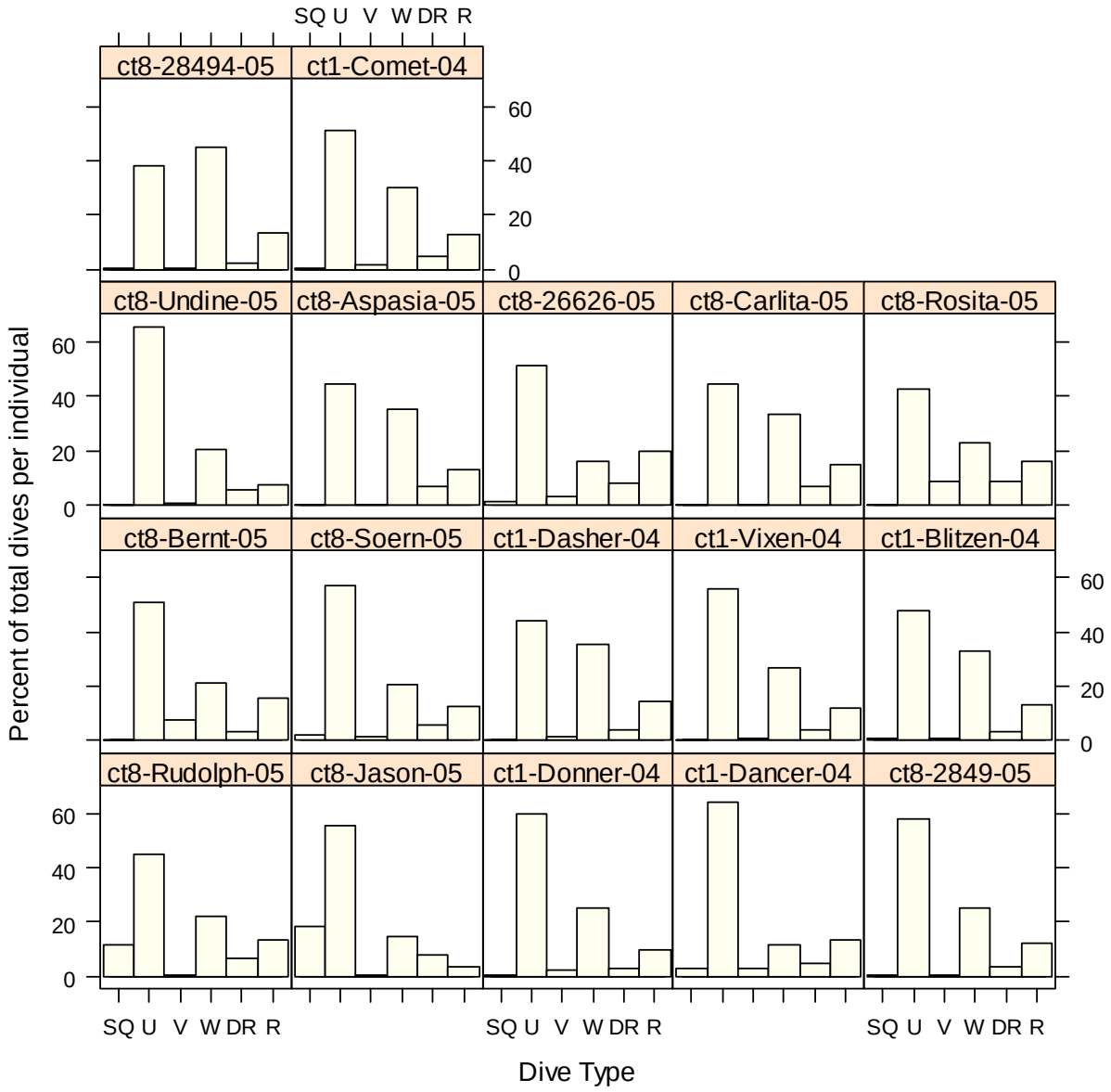

Figure 14. Relative occurrence of dive type (SQ: square dive, U: u-shaped dive, V: v-shaped dive, W: wiggle dive, DR: drift dive, R: root dive) in individual southern elephant seals.

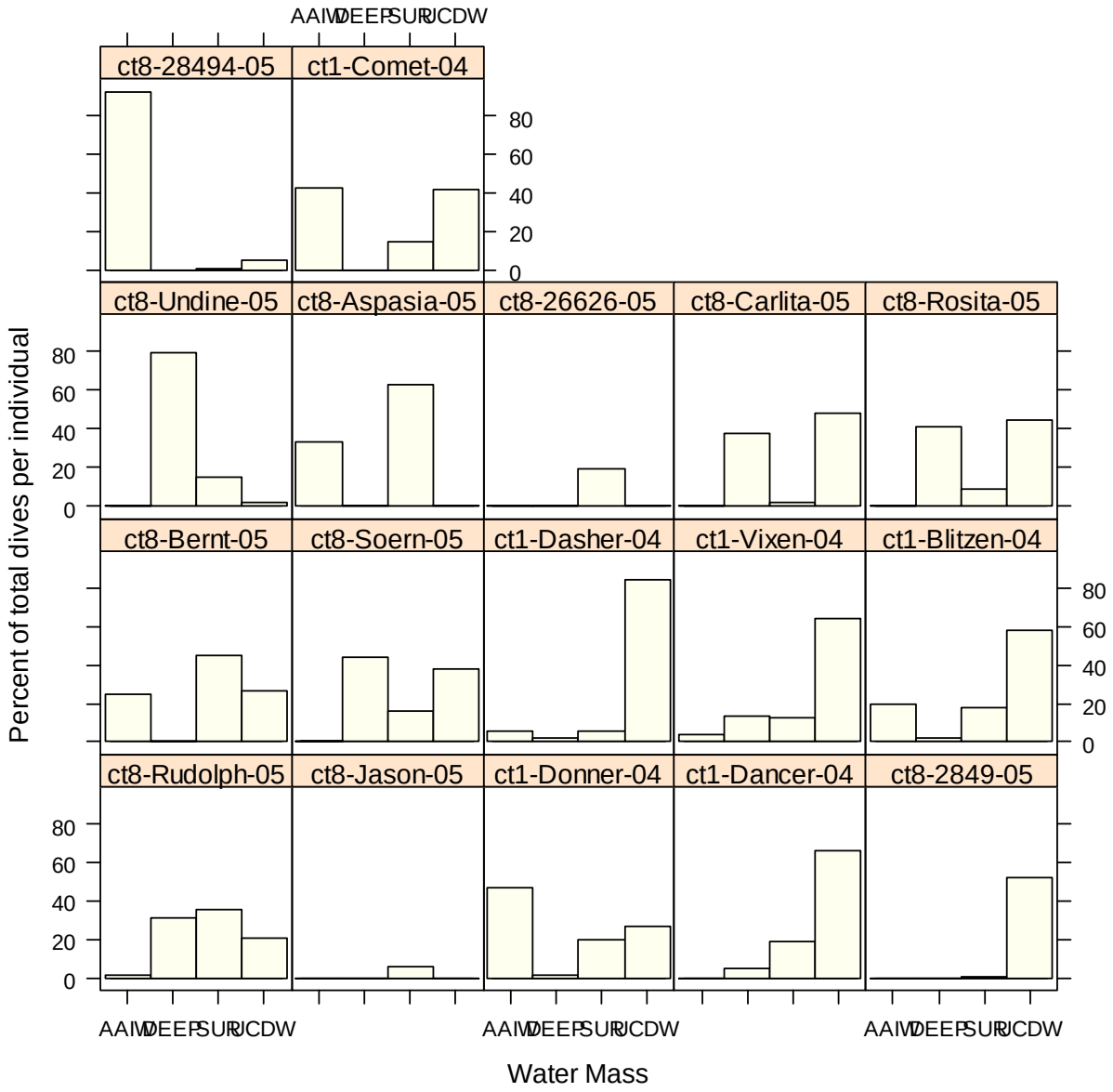

Figure 15. Water mass use in individual southern elephant seals (AAIW: Antarctic Intermediate Water, DEEP: Bottom Water, SUR: Surface Water, UCDW: Upper Circumpolar Deep Water)

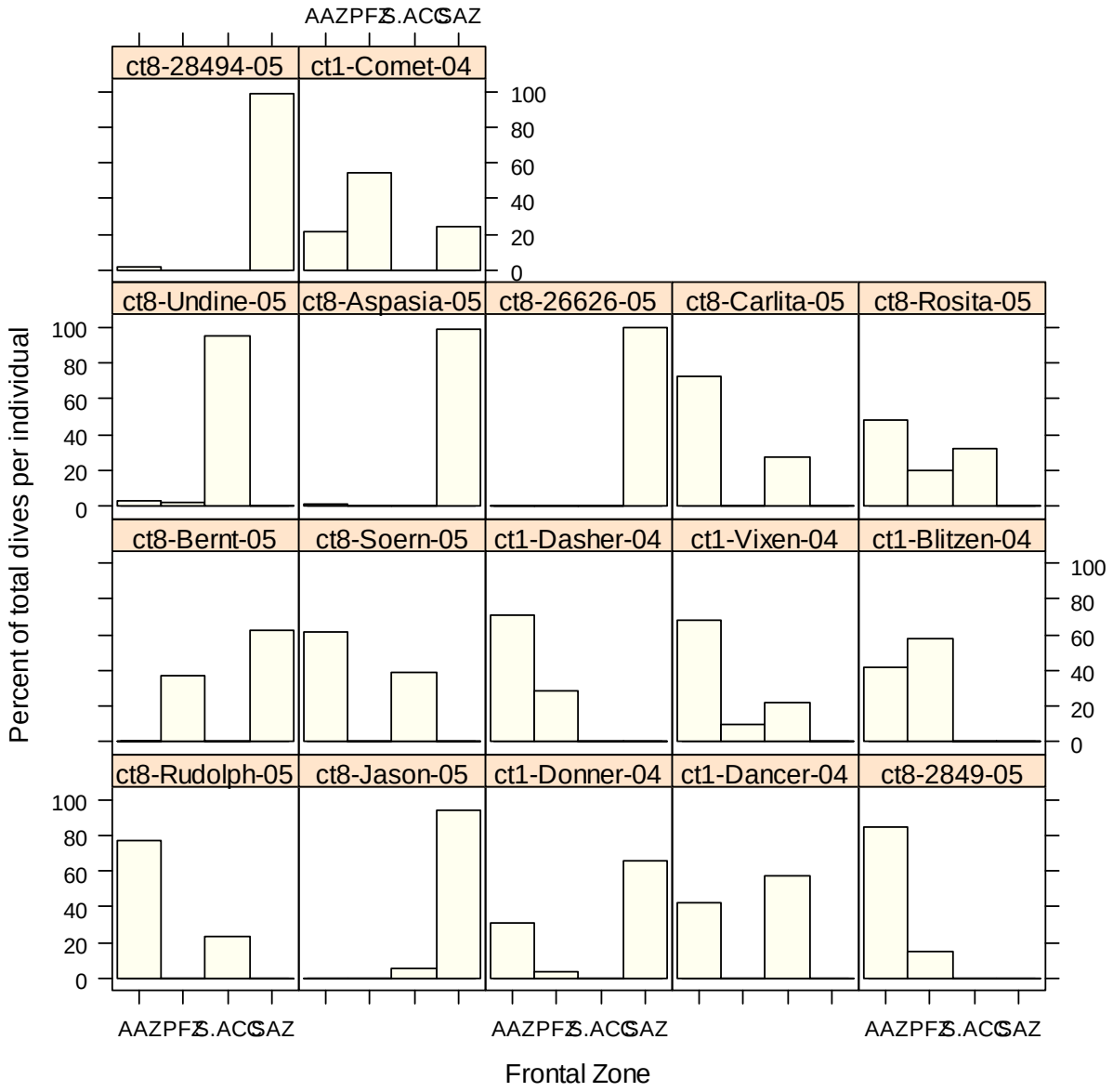

Figure 16. Frontal zone use in individual southern elephant seals (S.ACC: South of Antarctic Circumpolar Current, AAZ: Antarctic Current, PFZ: Polar Frontal Zone, SAZ: Subantarctic Zone)

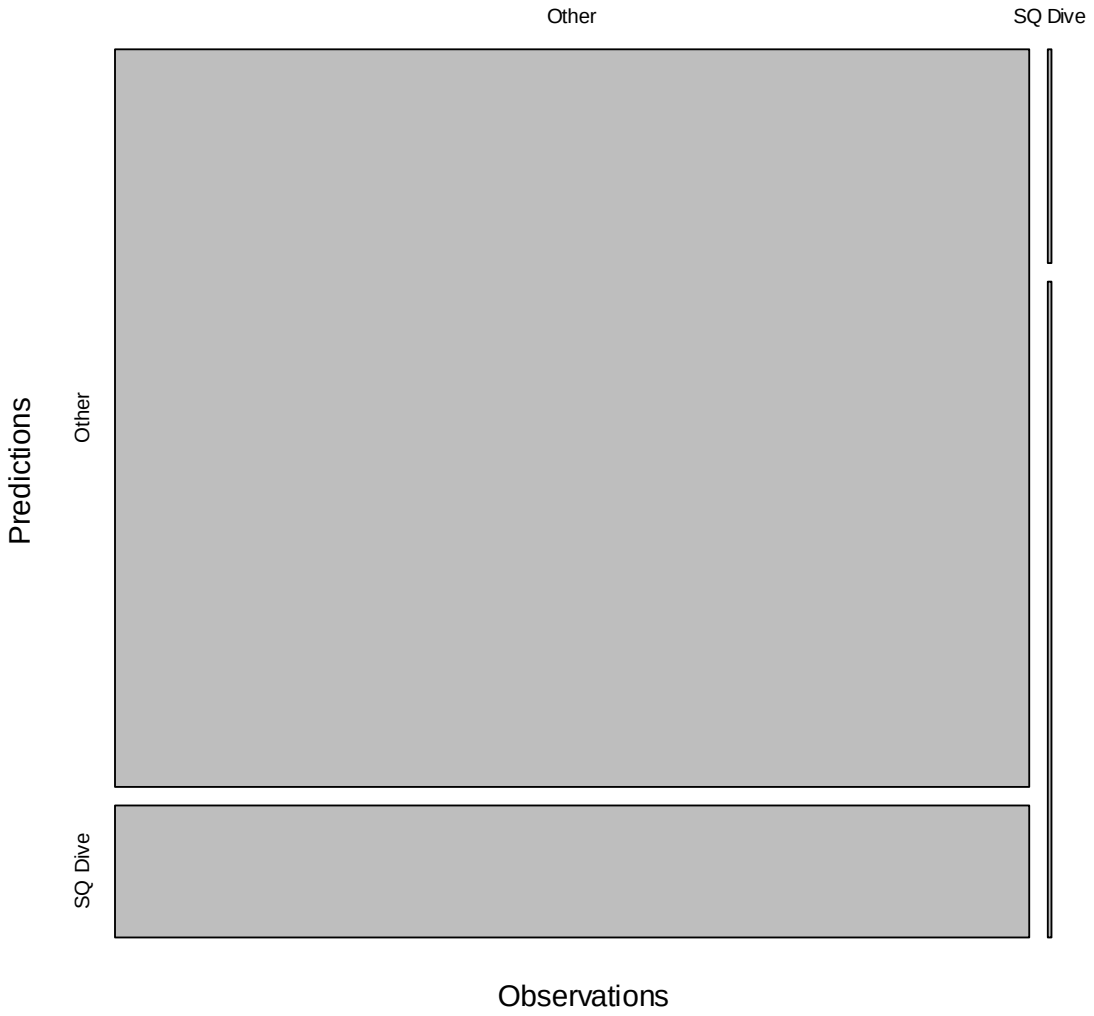

Figure 17. Representation of the confusions matrix of observed and predicted occurrences of SQ dives by GLM.

Table 8. Confusion matrix of predicted occurrences by GLM versus observations of SQ dives

	Observations	
Predictions	Other	SQ dive
Other	17768	19
SQ dive	3170	58

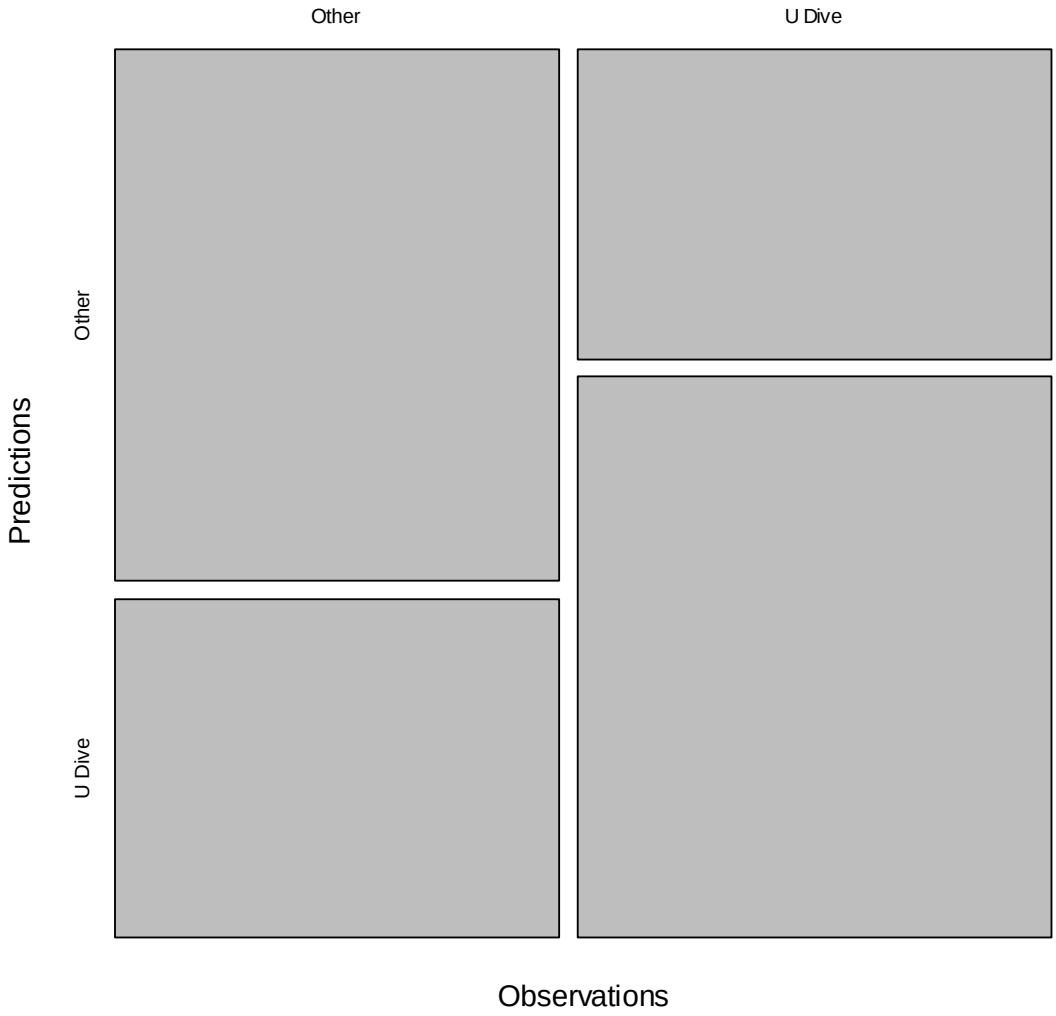

Figure 18. Representation of the confusions matrix of observed and predicted occurrences of U dives by GLM.

Table 9. Confusion matrix of predicted occurrences by GLM versus observations of U dives

	Observations	
Predictions	Other	U dive
Other	6223	3942
U dive	3873	6977

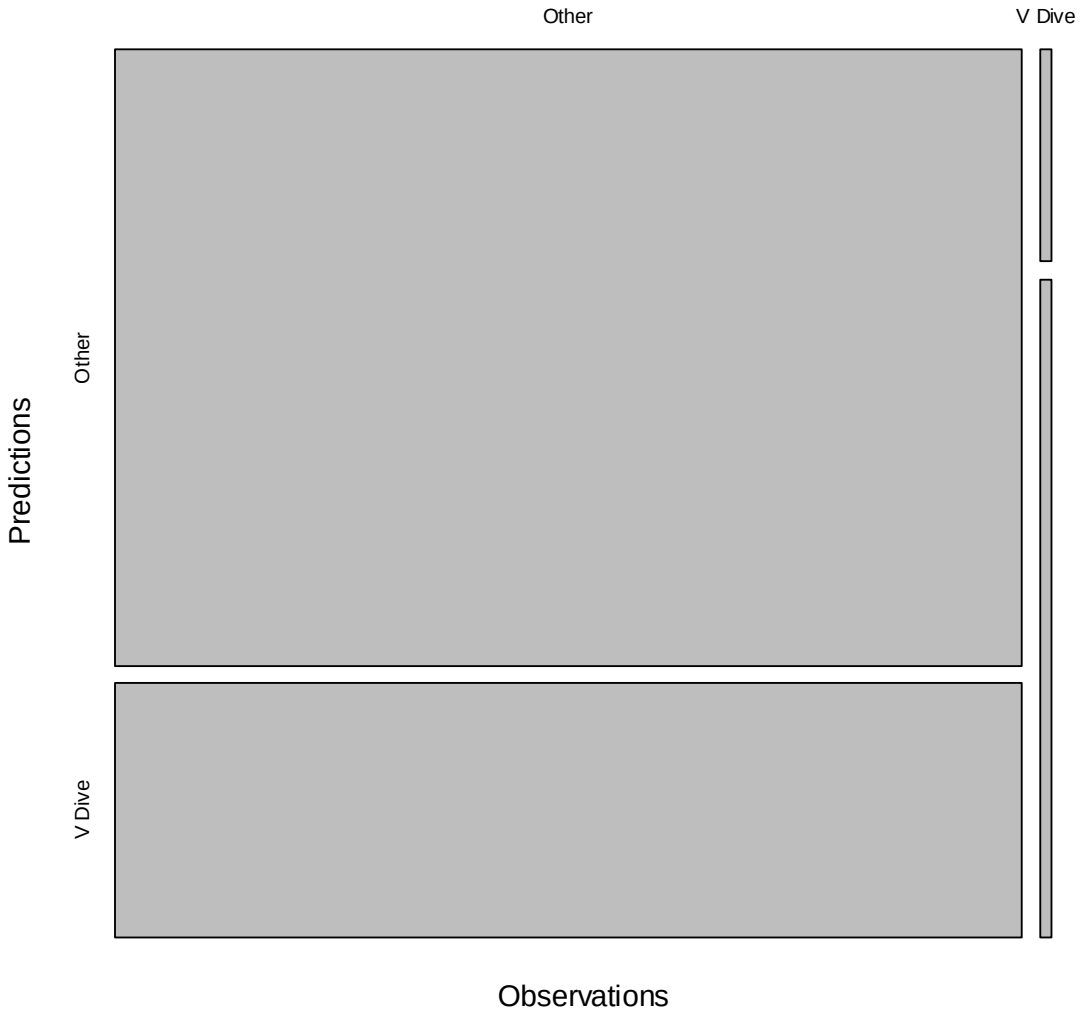

Figure 19. Representation of the confusions matrix of observed and predicted occurrences of V dives by GLM.

Table 10. Confusion matrix of predicted occurrences by GLM versus observations of V dives

	Observations	
Predictions	Other	V dive
Other	14724	6050
V dive	59	182

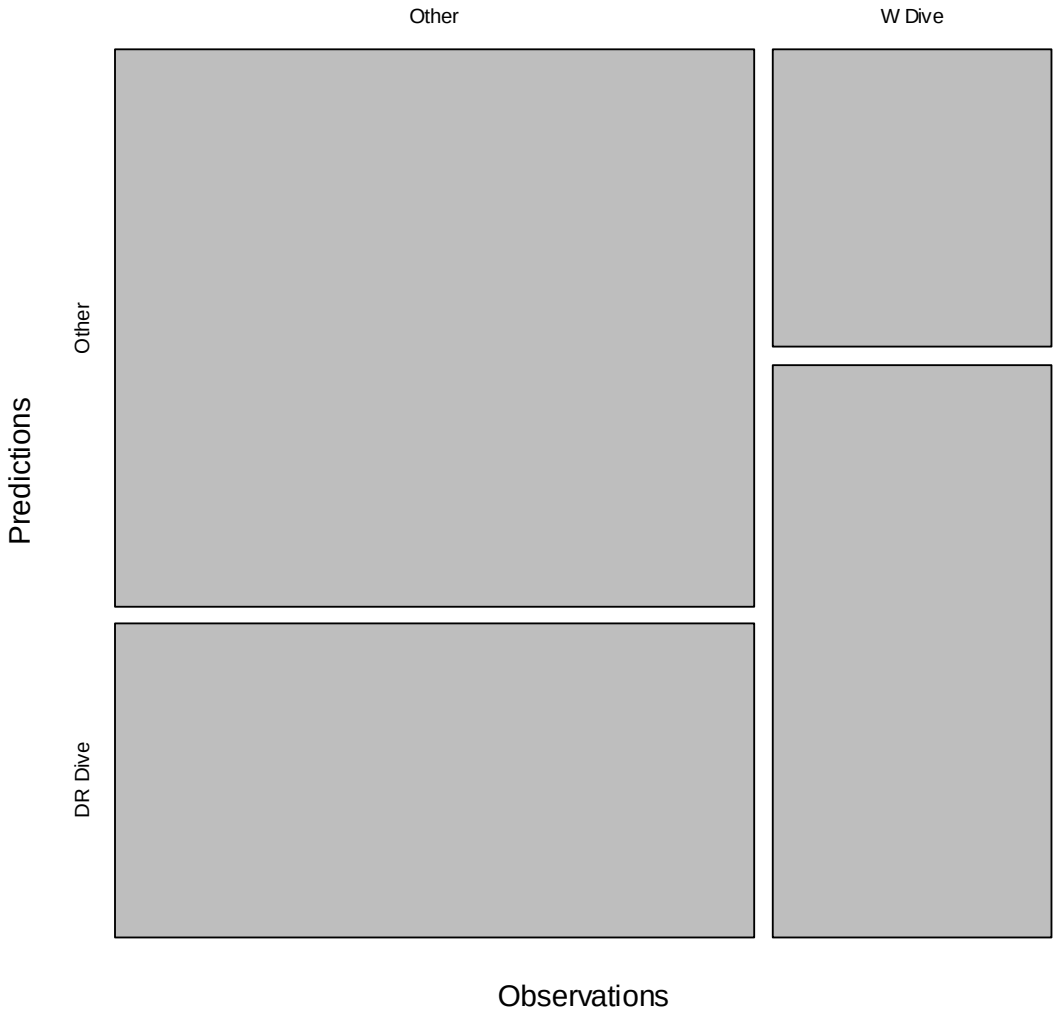

Figure 20. Representation of the confusions matrix of observed and predicted occurrences of W dives by GLM.

Table 11. Confusion matrix of predicted occurrences by GLM versus observations of W dives

	Observations	
Predictions	Other	W dive
Other	9377	5263
W dive	2184	4191

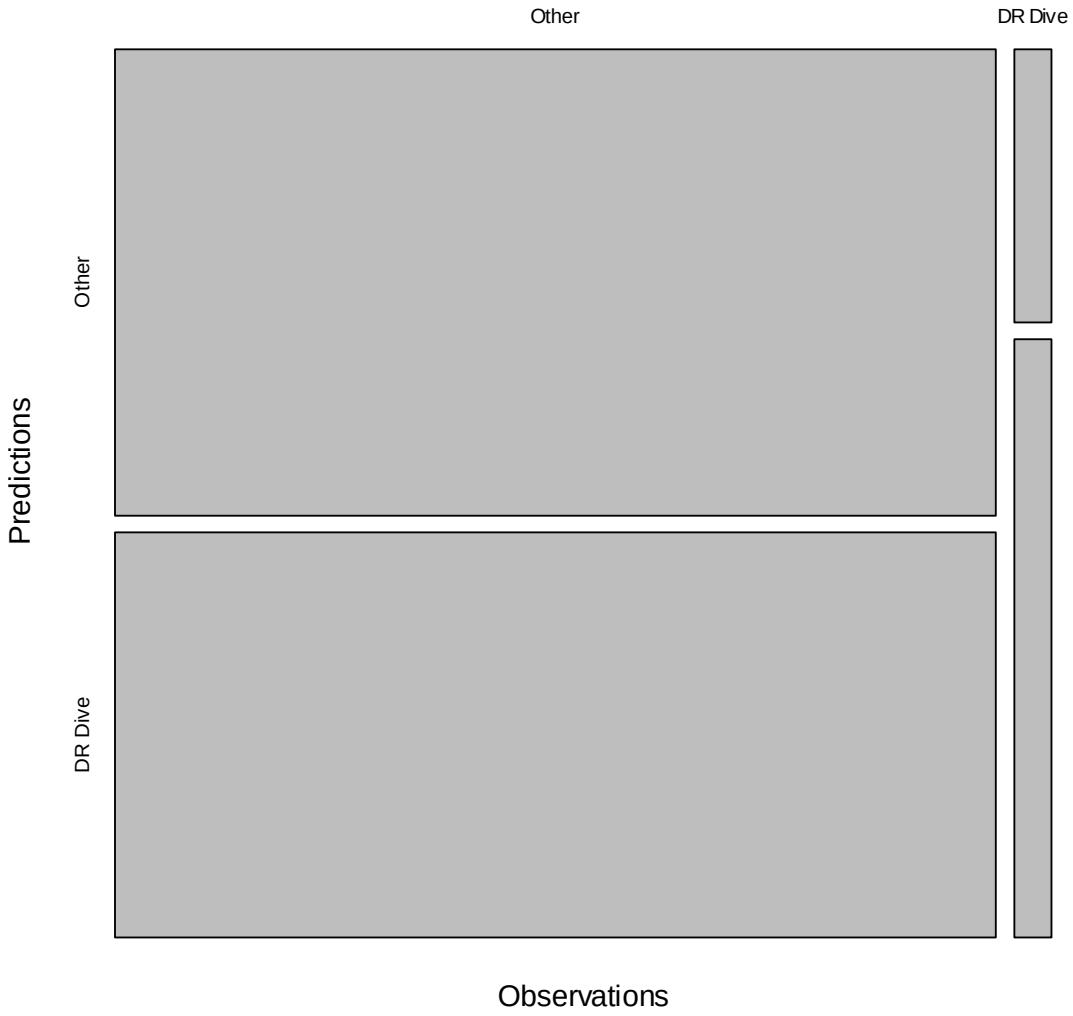

Figure 21. Representation of the confusions matrix of observed and predicted occurrences of DR dives by GLM.

Table 12. Confusion matrix of predicted occurrences by GLM versus observations of DR dives

	Observations	
Predictions	Other	DR dive
Other	10808	9366
DR dive	264	577

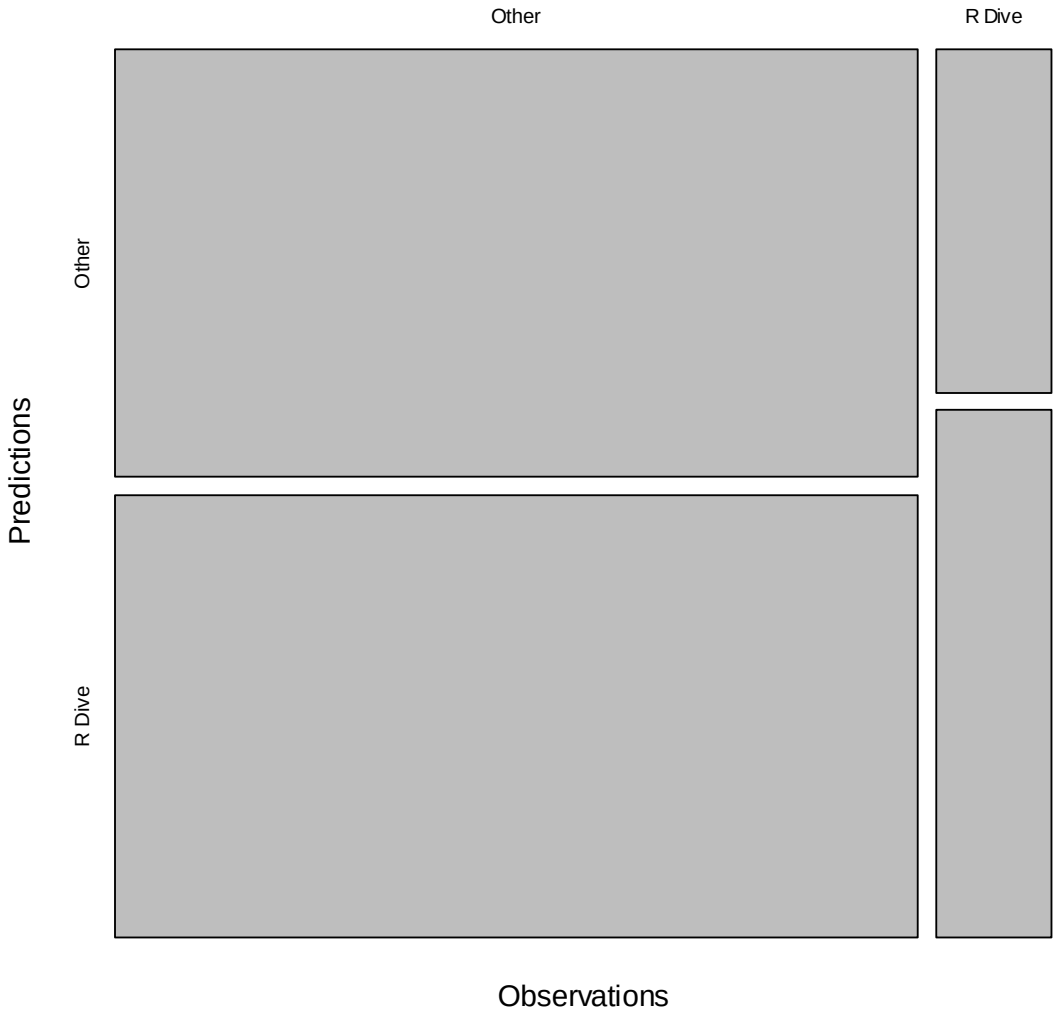

Figure 22. Representation of the confusions matrix of observed and predicted occurrences of R dives by GLM.

Table 13. Confusion matrix of predicted occurrences by GLM versus observations of R dives

	Observations	
Predictions	Other	R dive
Other	9046	9338
R dive	1039	1592

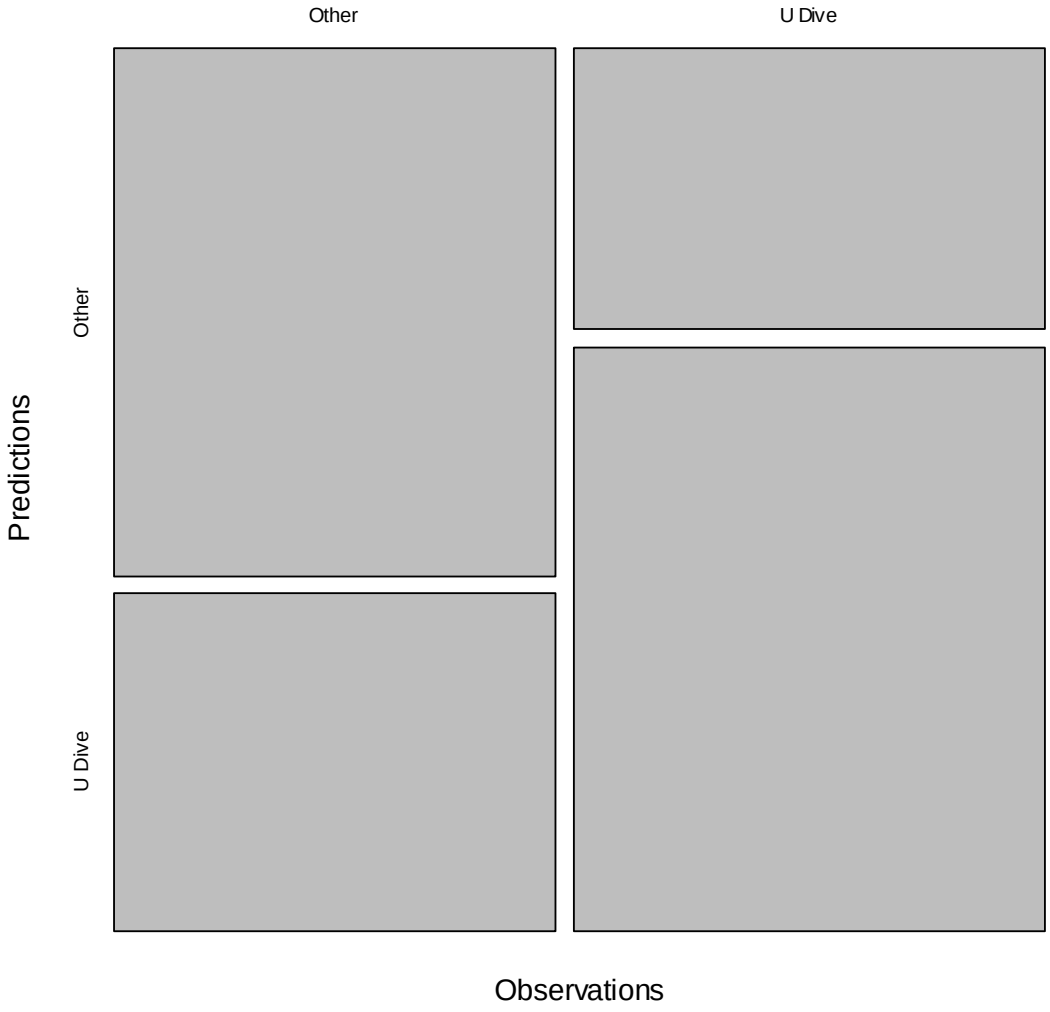

Figure 23. Representation of the confusions matrix of observed and predicted occurrences of U dives by GAM.

Table 14. Confusion matrix of predicted occurrences by GAM versus observations of U dives

	Observations	
Predictions	Other	U dive
Other	6205	3960
U dive	3529	7321

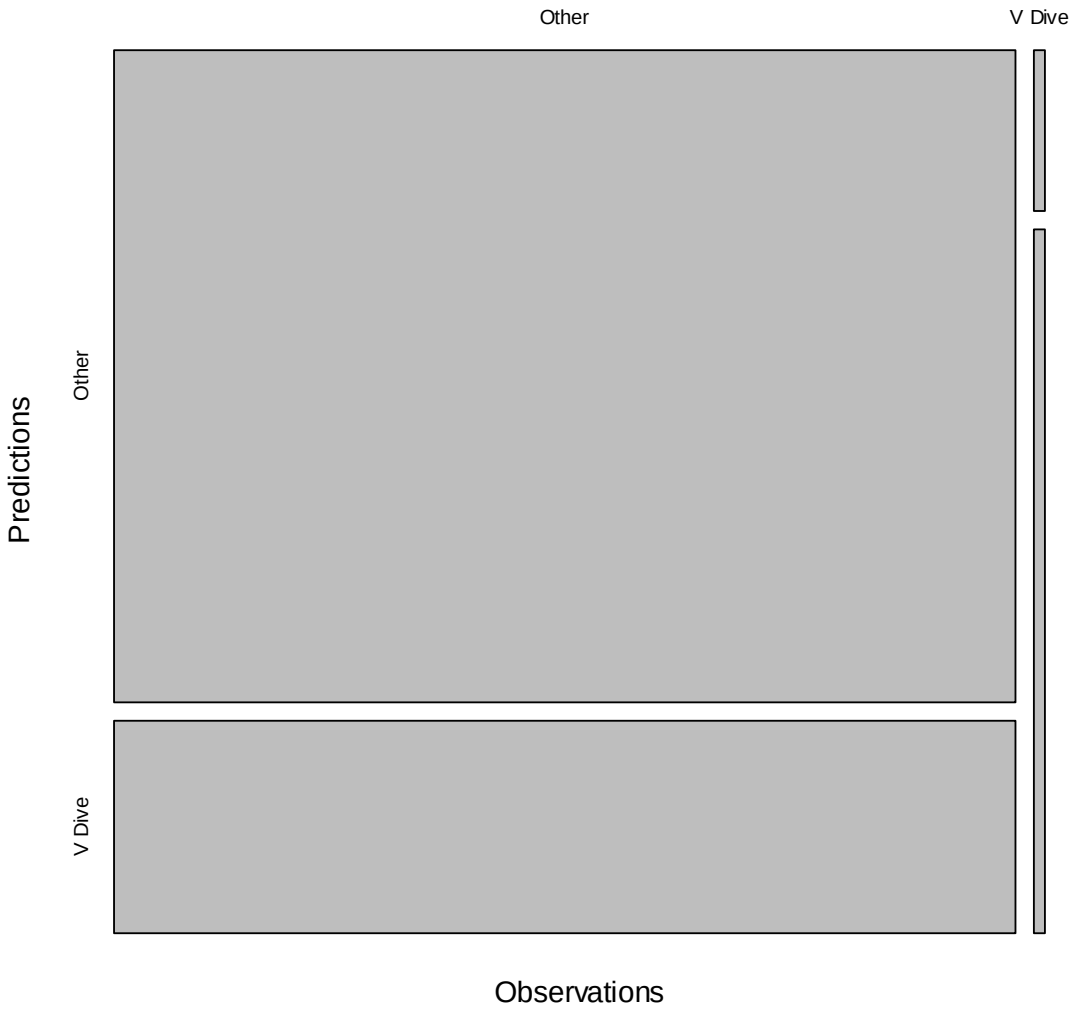

Figure 24. Representation of the confusions matrix of observed and predicted occurrences of V dives by GAM.

Table 15. Confusion matrix of predicted occurrences by GAM versus observations of V dives

	Observations	
Predictions	Other	V dive
Other	15675	5099
V dive	45	196

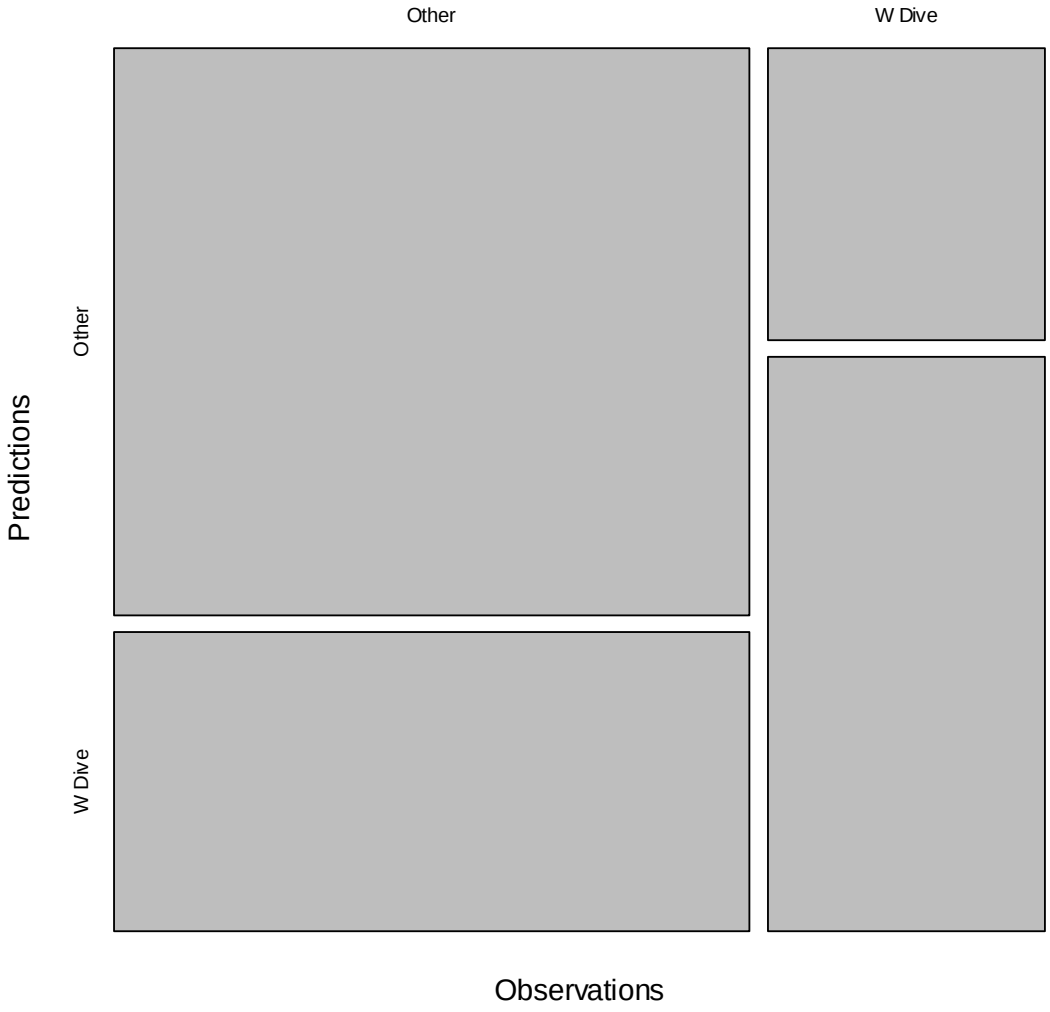

Figure 25. Representation of the confusions matrix of observed and predicted occurrences of W dives by GAM.

Table 16. Confusion matrix of predicted occurrences by GAM versus observations of W dives

	Observations	
Predictions	Other	W dive
Other	9596	5044
W dive	2153	4222

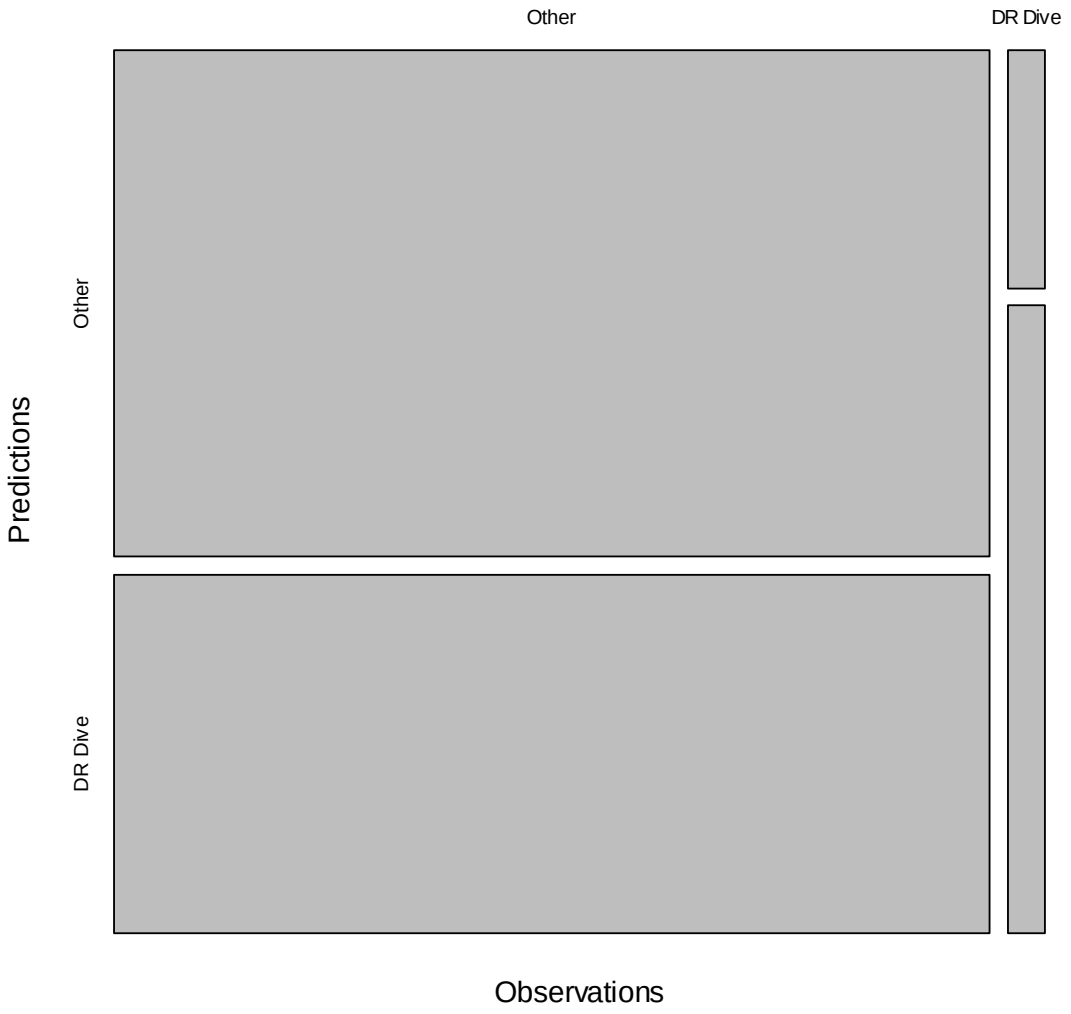

Figure 26. Representation of the confusions matrix of observed and predicted occurrences of DR dives by GAM.

Table 17. Confusion matrix of predicted occurrences by GAM versus observations of DR dives

	Observations	
Predictions	Other	DR dive
Other	11814	8360
DR dive	232	609

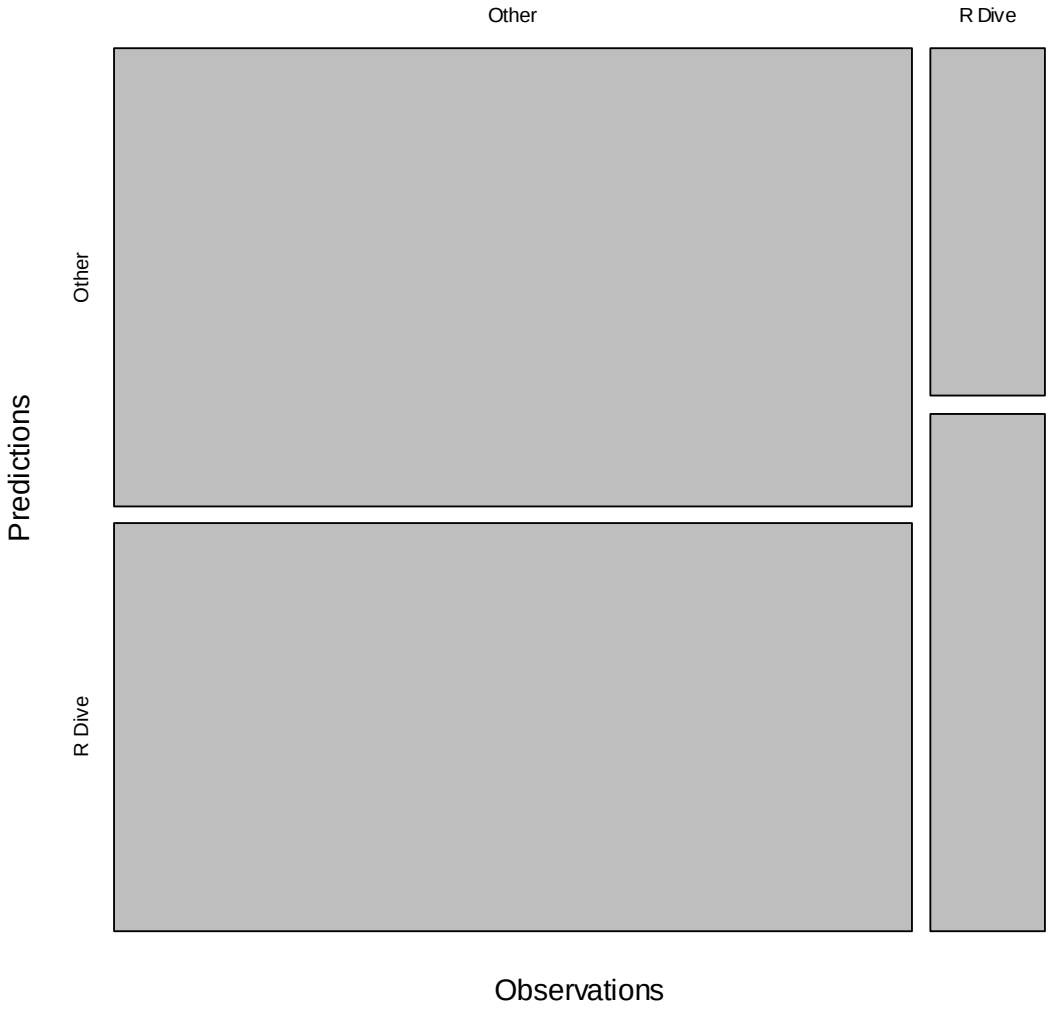

Figure 27. Representation of the confusions matrix of observed and predicted occurrences of R dives by GAM.

Table 18. Confusion matrix of predicted occurrences by GAM versus observations of R dives

	Observations	
Predictions	Other	R dive
Other	9730	8654
R dive	1060	1571

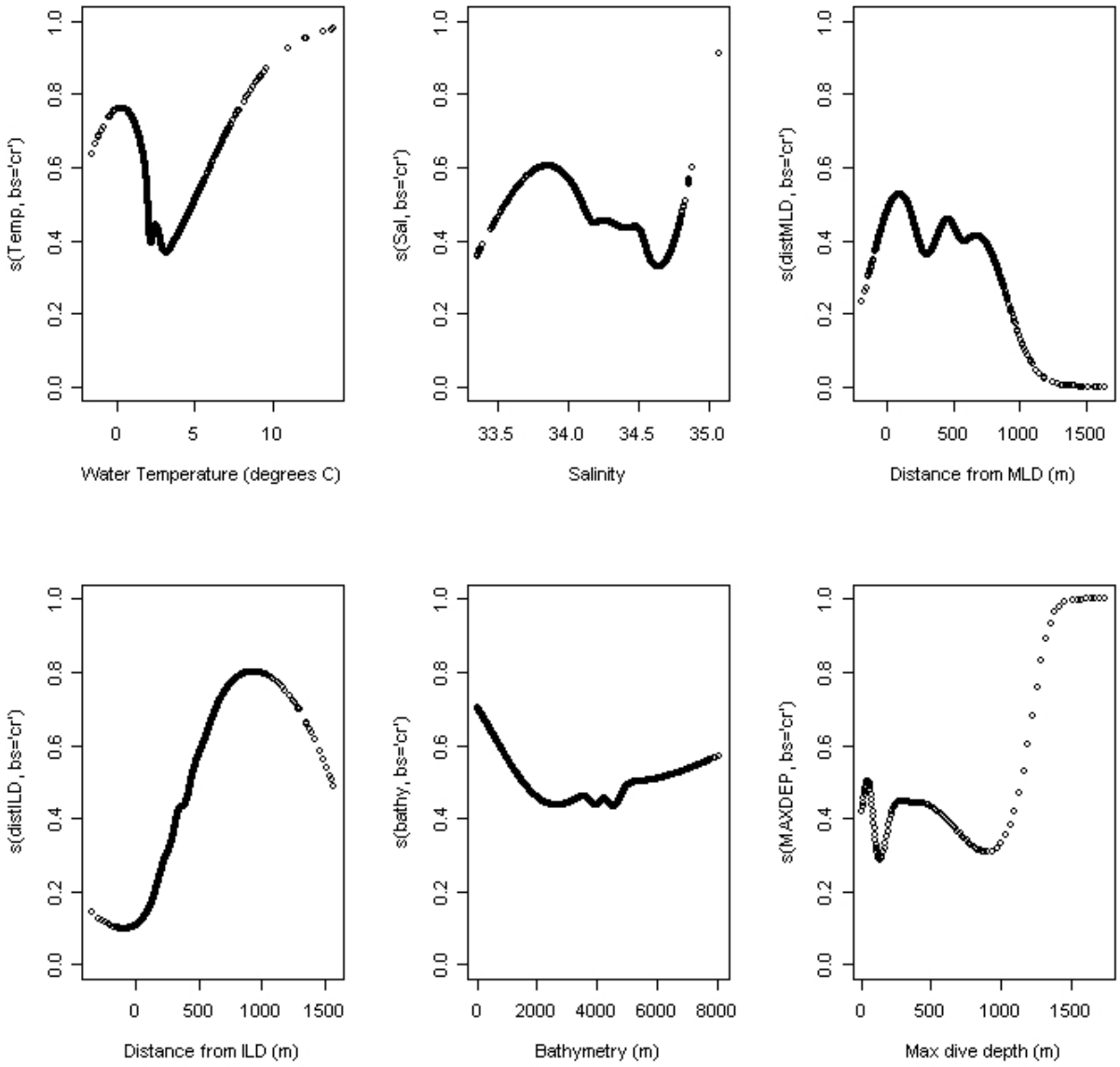

Figure 28. Smooth functions on the scale of the response for the GAM fitted to binary data for U dives.

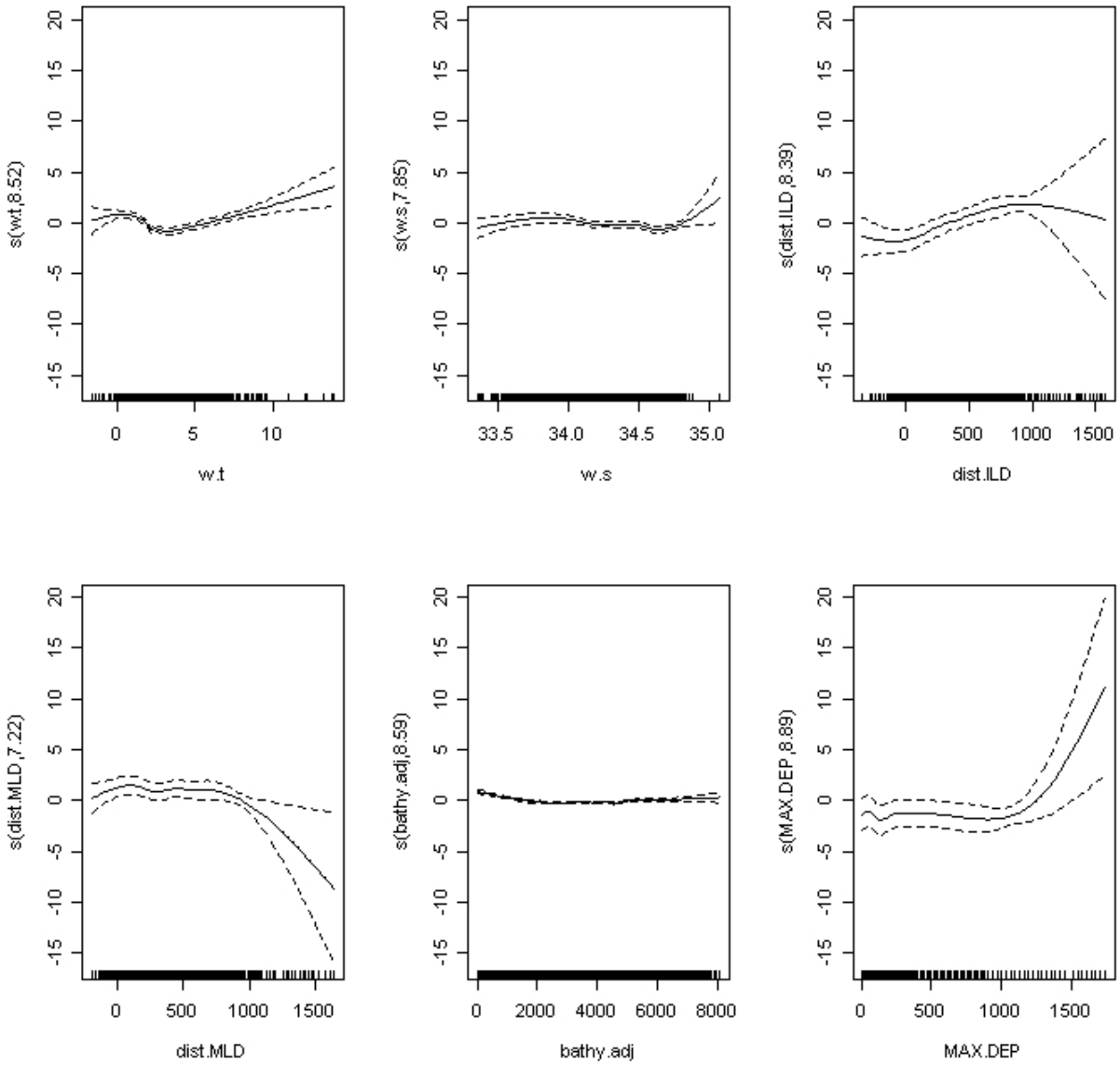

Figure 29. Smooth functions on the scale of the linear predictor for the GAM fitted to binary data for U dives (w.t – Temperature degrees ºC, w.s - Salinity, dist.MLD - Distance from MLD (m), dist.ILD - Distance from ILD (m), MAX.DEP – Maximum dive depth (m), bathy.adj – Bathymetry (m))

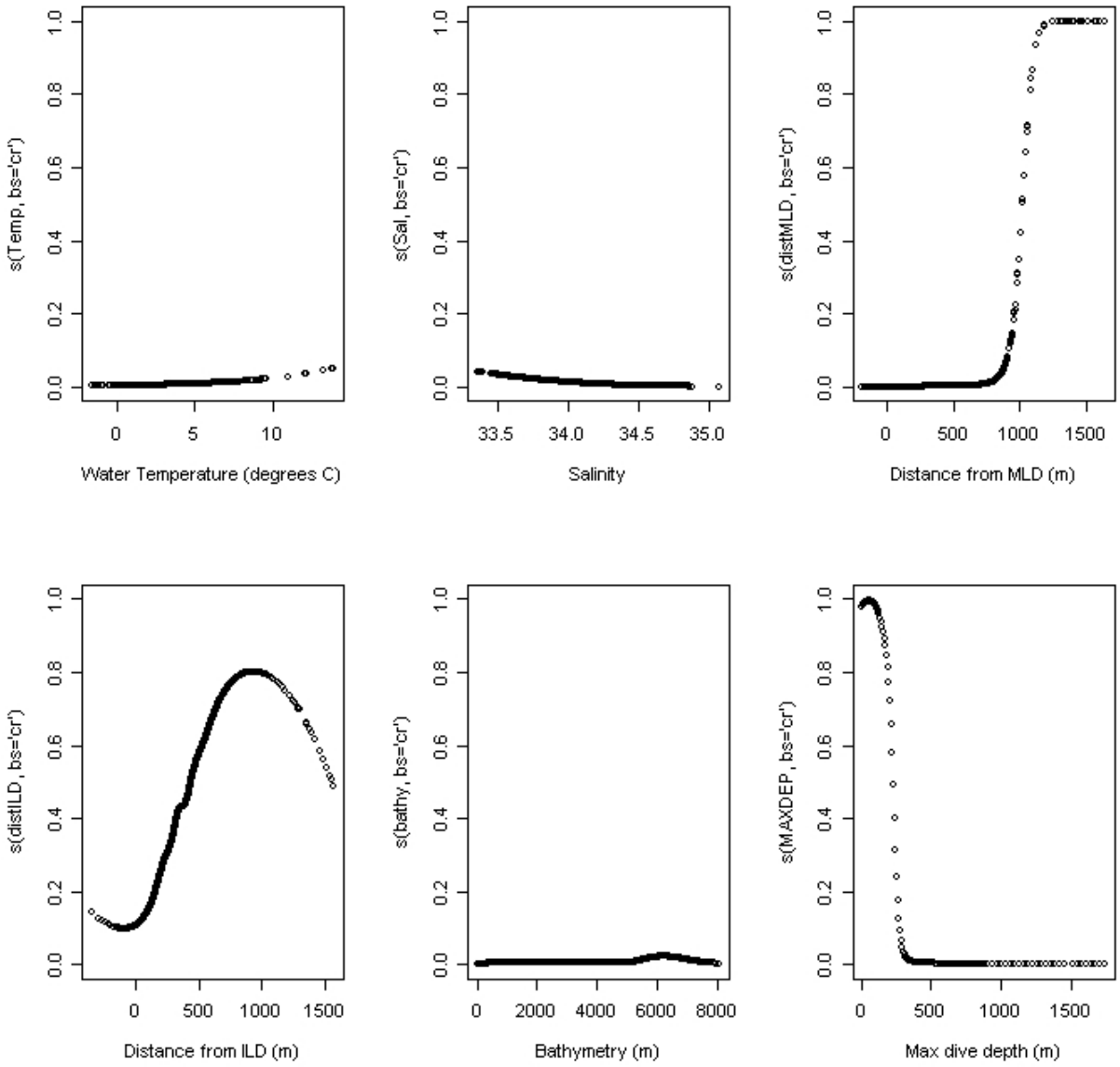

Figure 30. Smooth functions on the scale of the response for the GAM fitted to binary data for V dives.

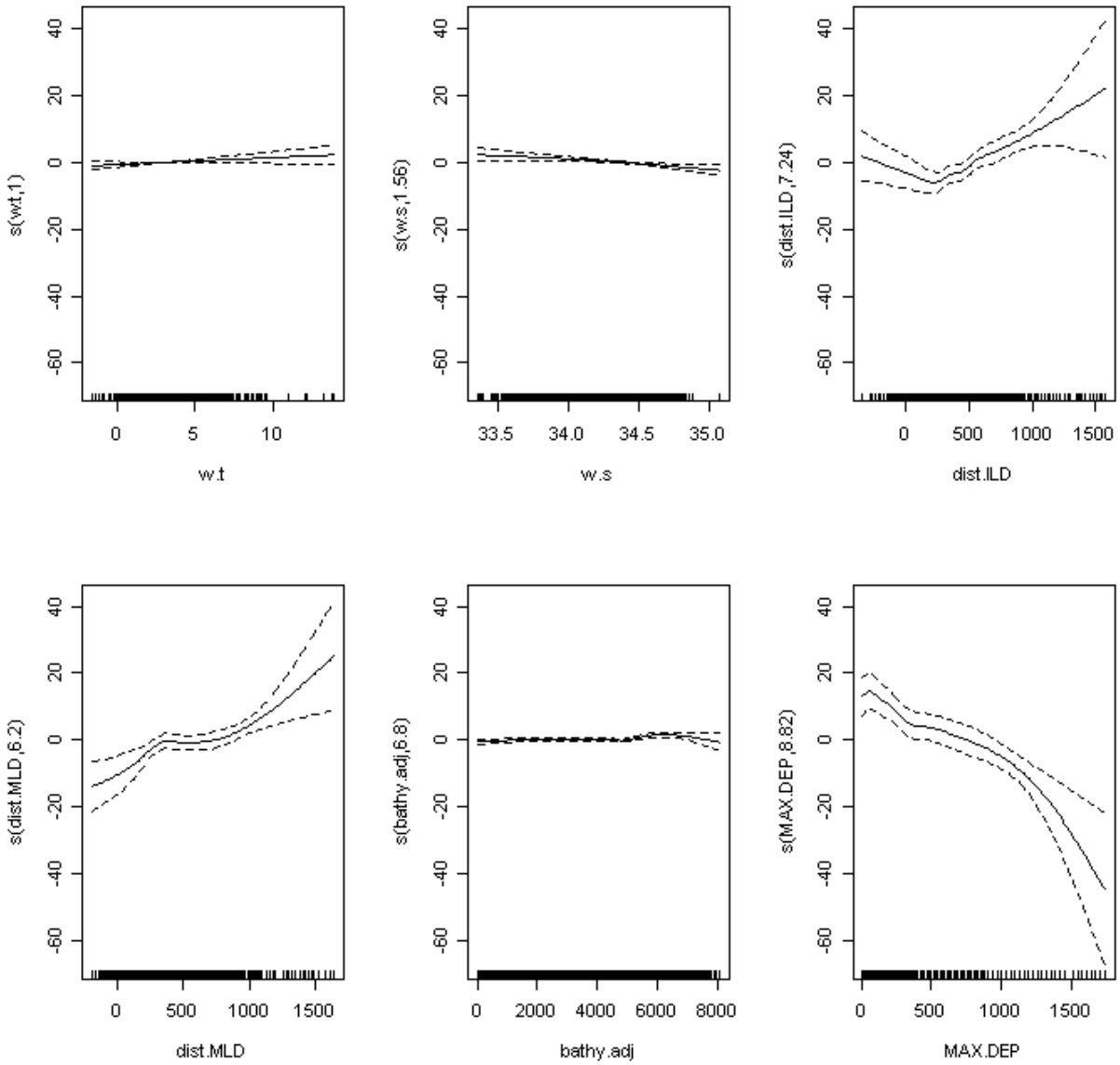

Figure 31. Smooth functions on the scale of the linear predictor for the GAM fitted to binary data for V dives (w.t – Temperature degrees ºC, w.s - Salinity, dist.MLD - Distance from MLD (m), dist.ILD - Distance from ILD (m), MAX.DEP – Maximum dive depth (m), bathy.adj – Bathymetry (m))

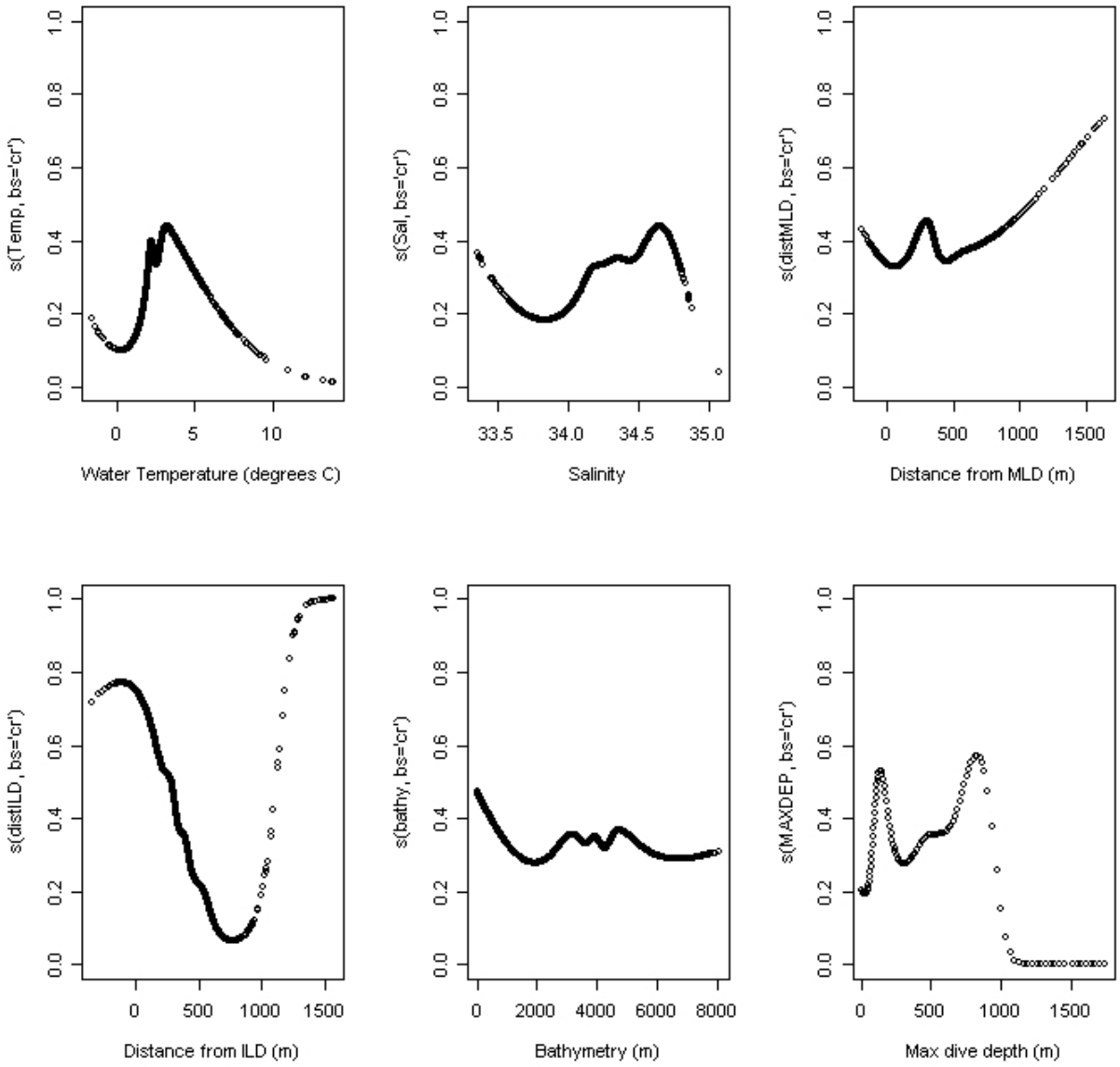

Figure 32. Smooth functions on the scale of the response for the GAM fitted to binary data for W dives.

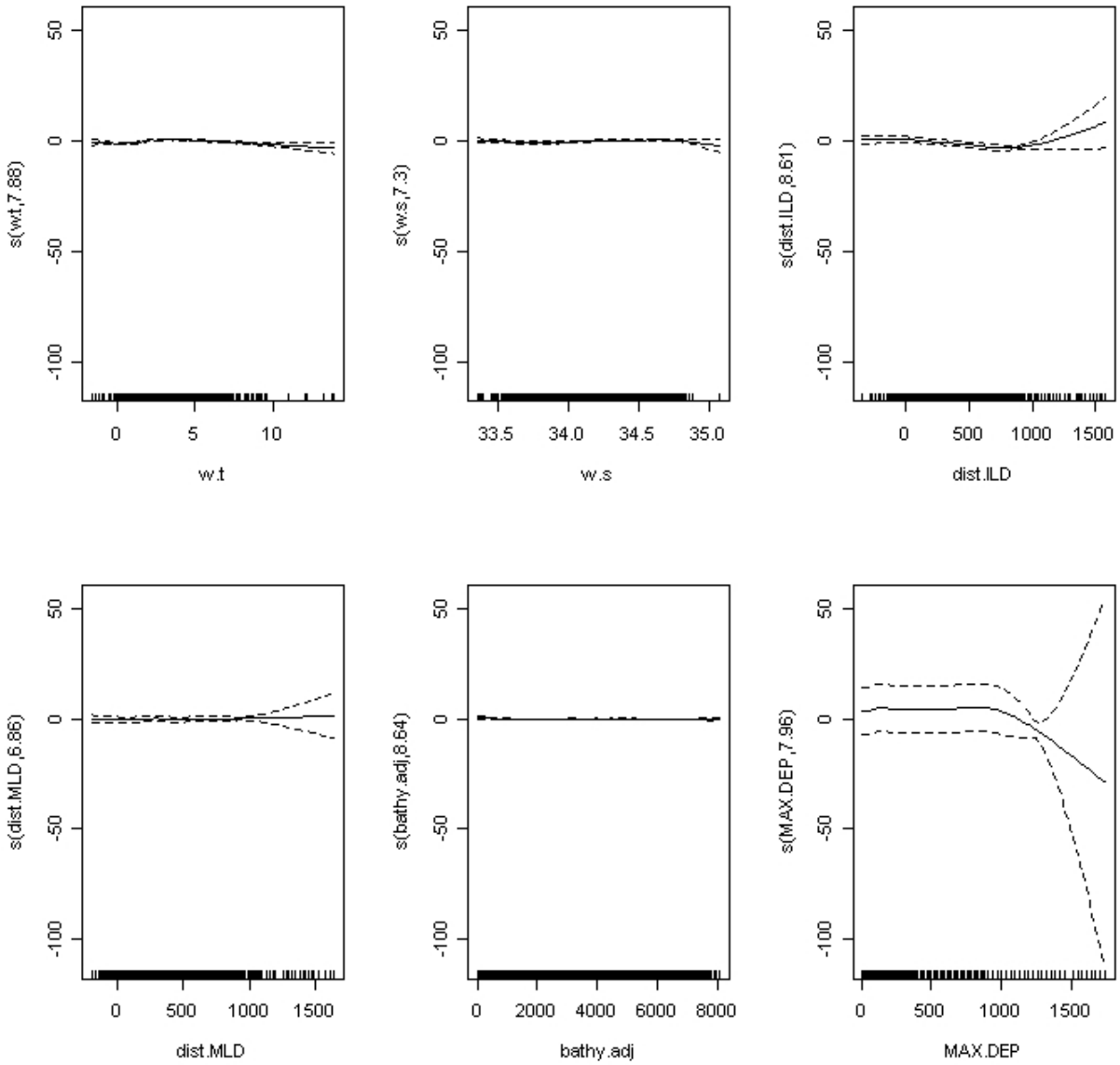

Figure 33. Smooth functions on the scale of the linear predictor for the GAM fitted to binary data for W dives (w.t – Temperature degrees ºC, w.s - Salinity, dist.MLD - Distance from MLD (m), dist.ILD - Distance from ILD (m), MAX.DEP – Maximum dive depth (m), bathy.adj – Bathymetry (m))

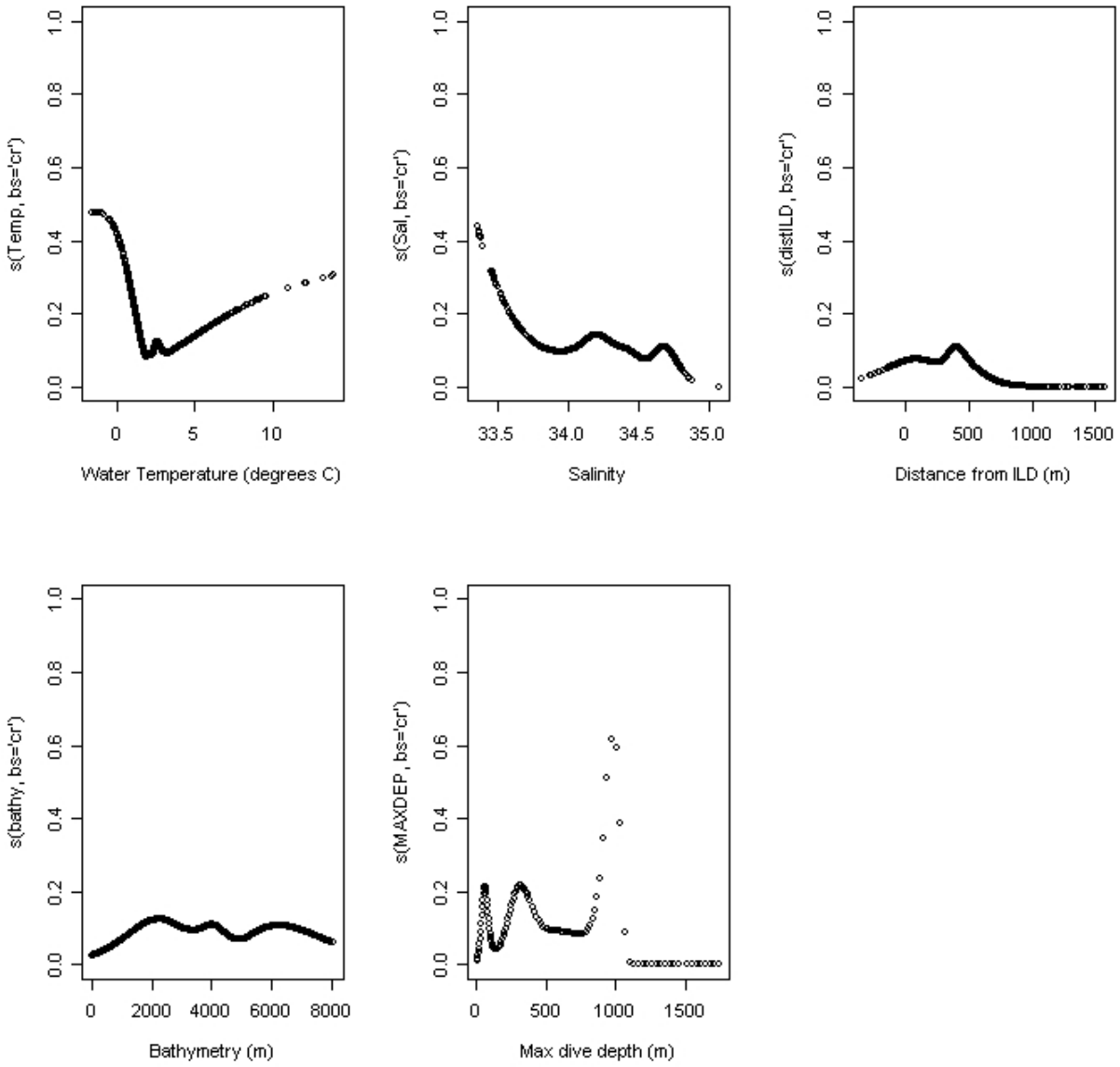

Figure 34. Smooth functions on the scale of the response for the GAM fitted to binary data for DR dives.

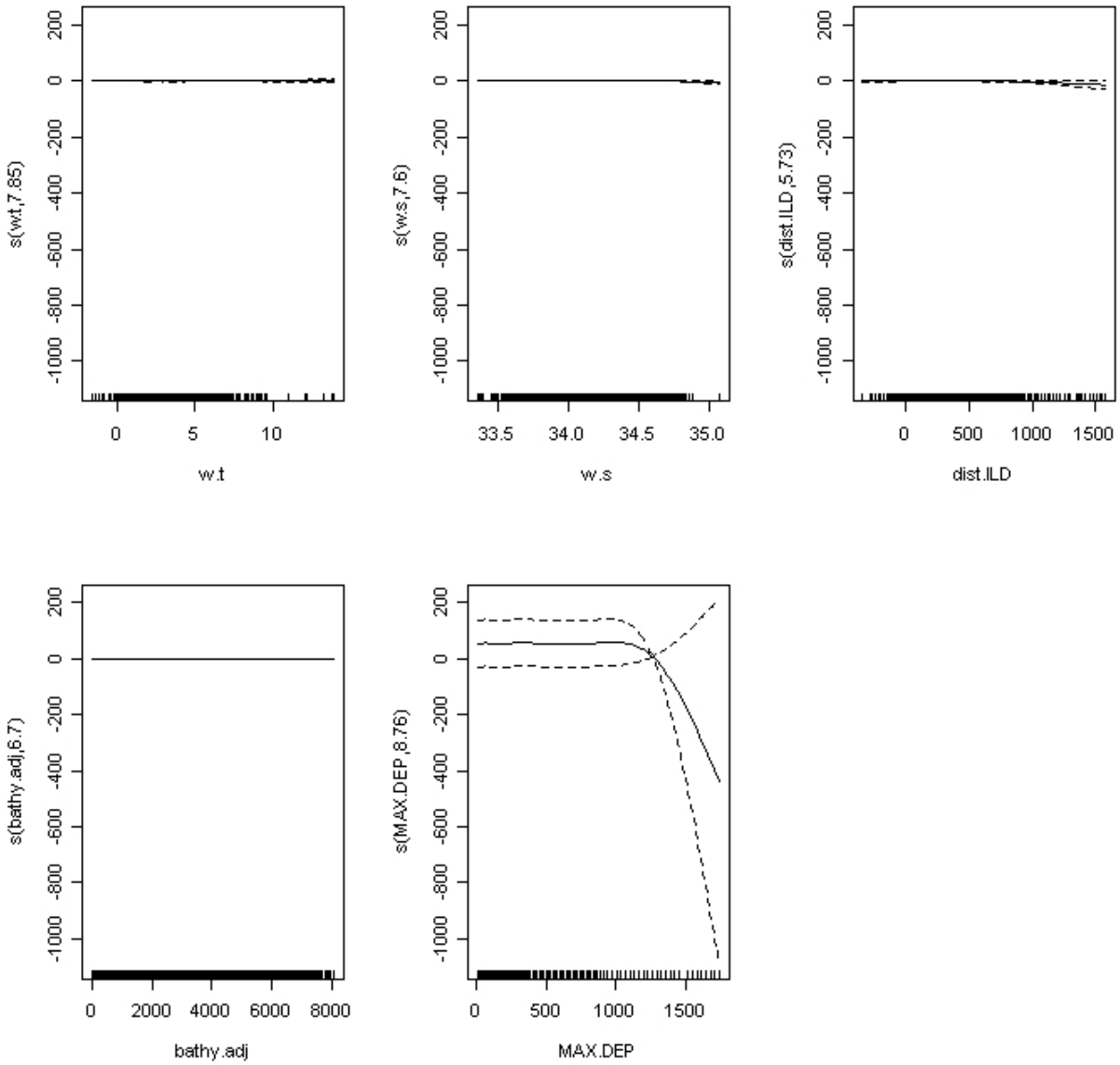

Figure 35. Smooth functions on the scale of the linear predictor for the GAM fitted to binary data for DR dives (w.t – Temperature degrees ºC, w.s - Salinity, dist.ILD - Distance from ILD (m), MAX.DEP – Maximum dive depth (m), bathy.adj – Bathymetry (m))

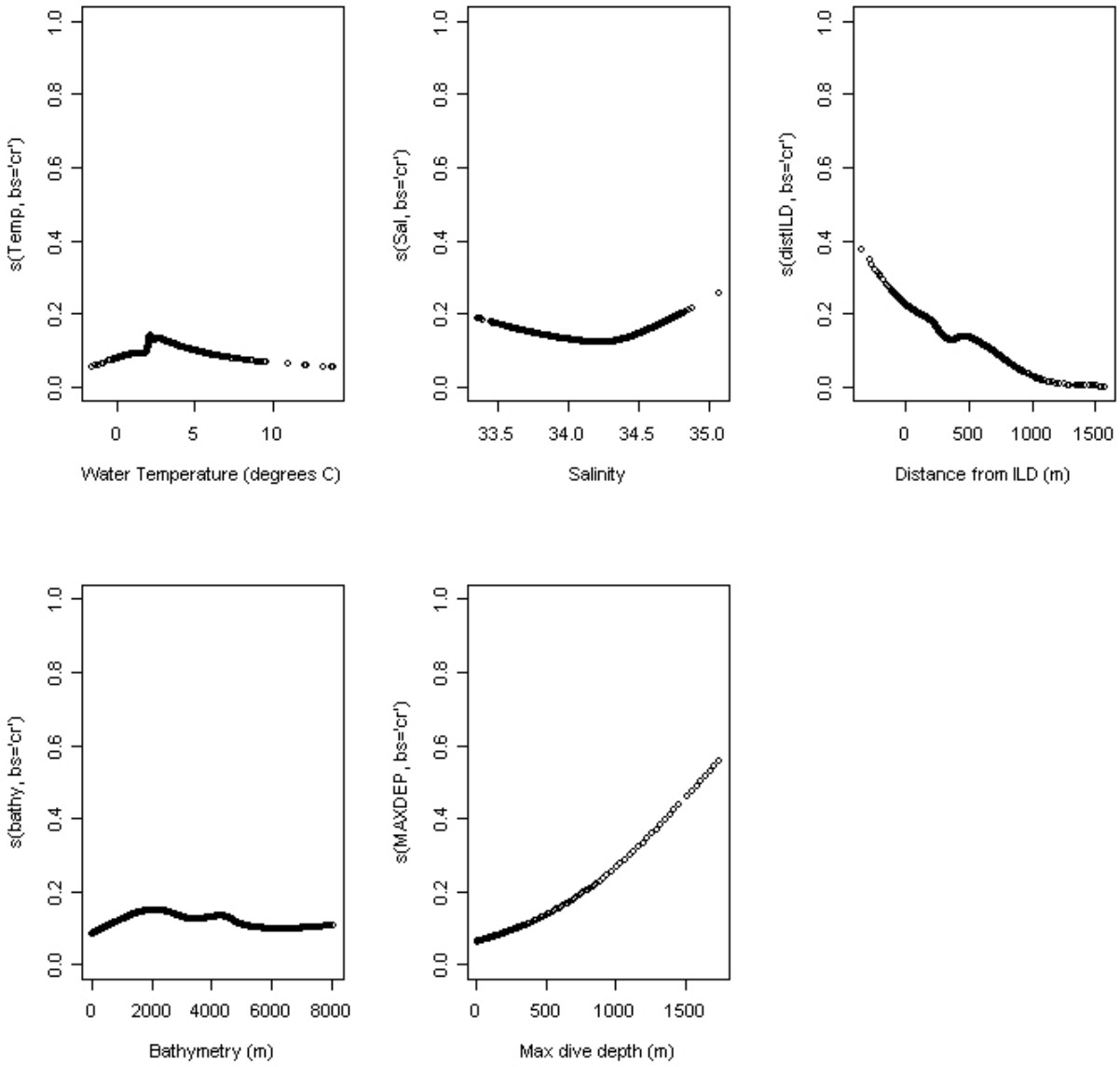

Figure 36. Smooth functions on the scale of the response for the GAM fitted to binary data for R dives.

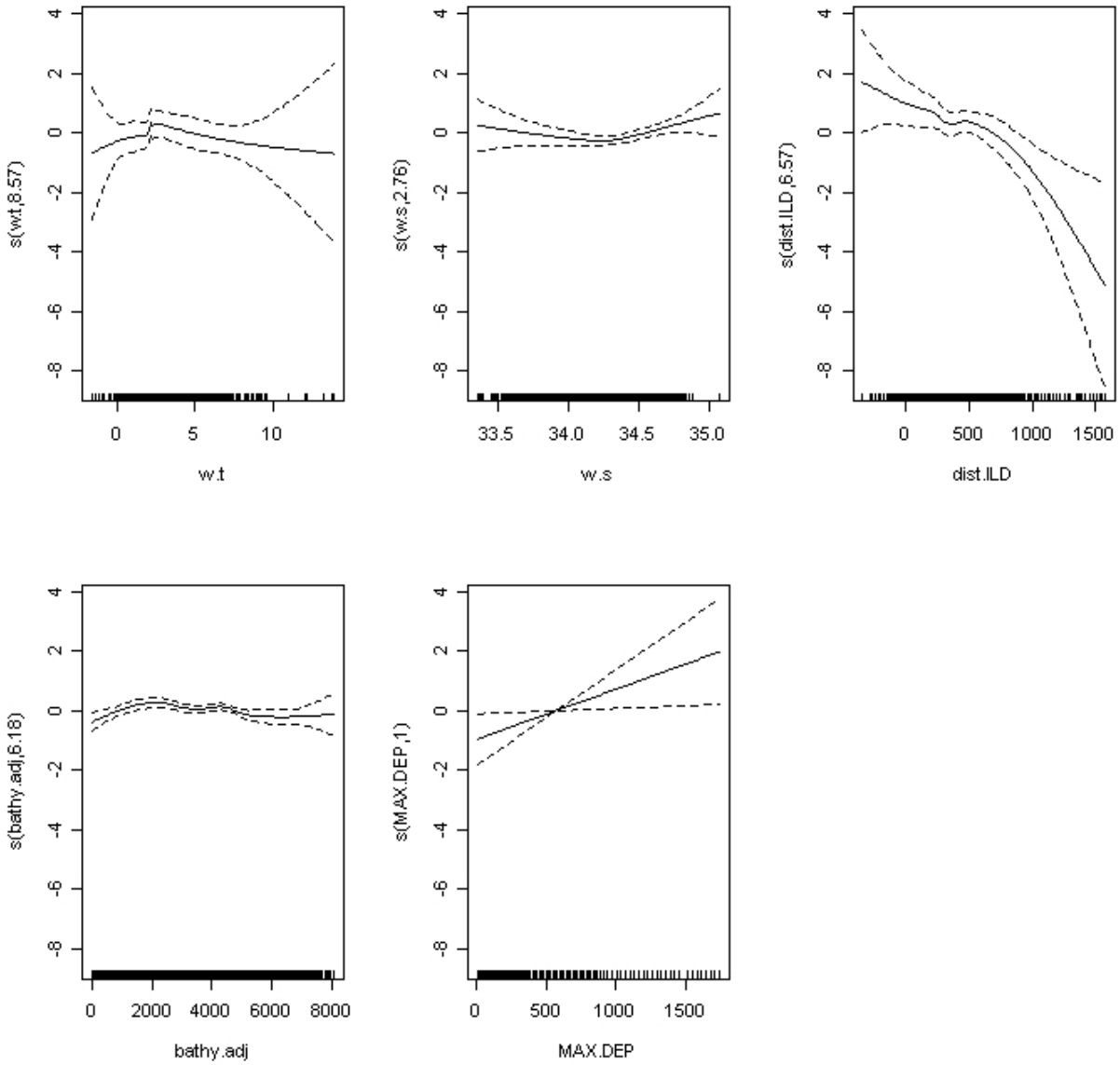

Figure 37. Smooth functions on the scale of the linear predictor for the GAM fitted to binary data for R dives (w.t – Temperature degrees ºC, w.s - Salinity, dist.ILD - Distance from ILD (m), MAX.DEP – Maximum dive depth (m), bathy.adj – Bathymetry (m))

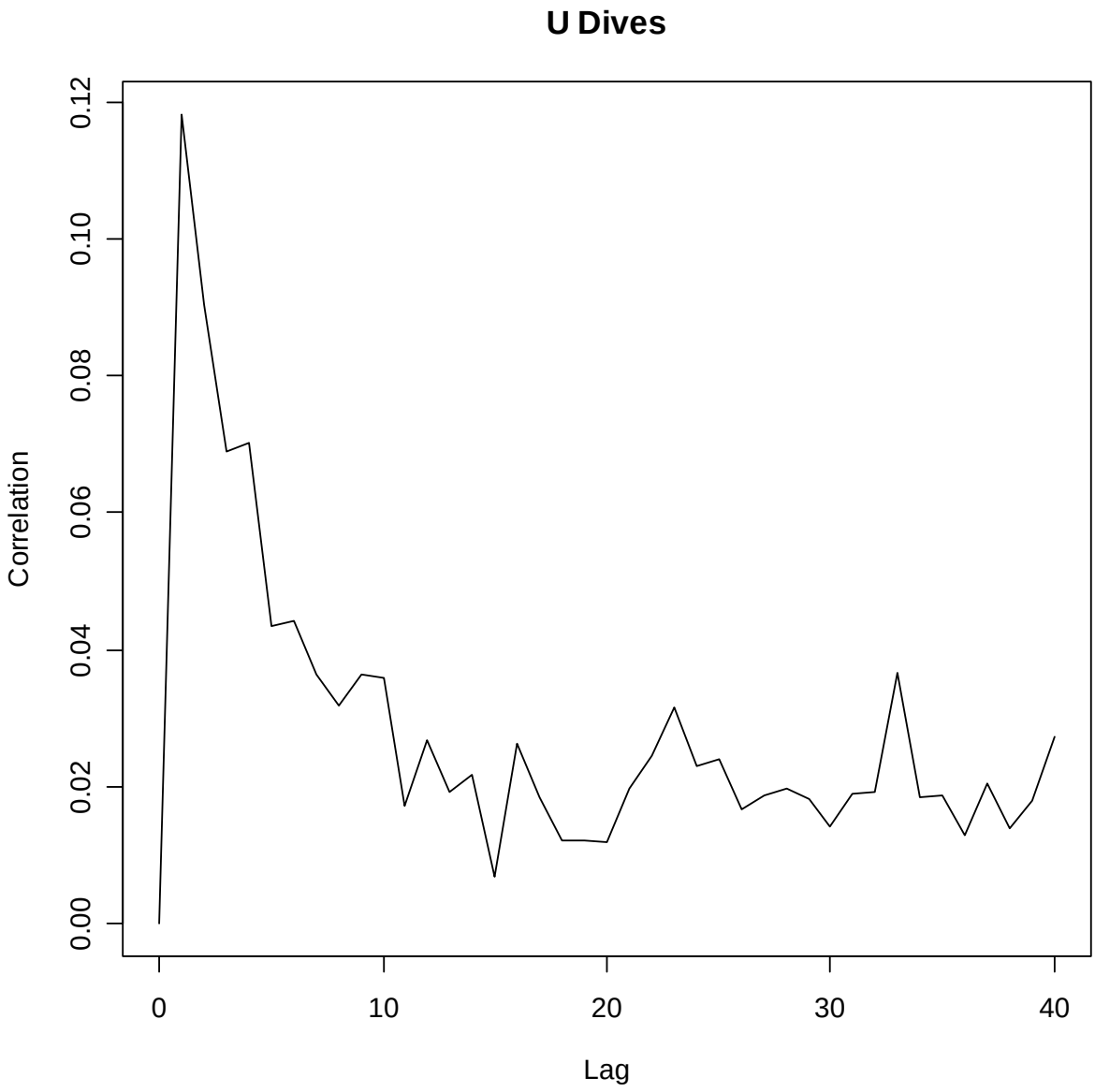

Figure 38. Correlogram of Pearson's residuals versus lag in dives for the GAM fitted to binary data for Udives.

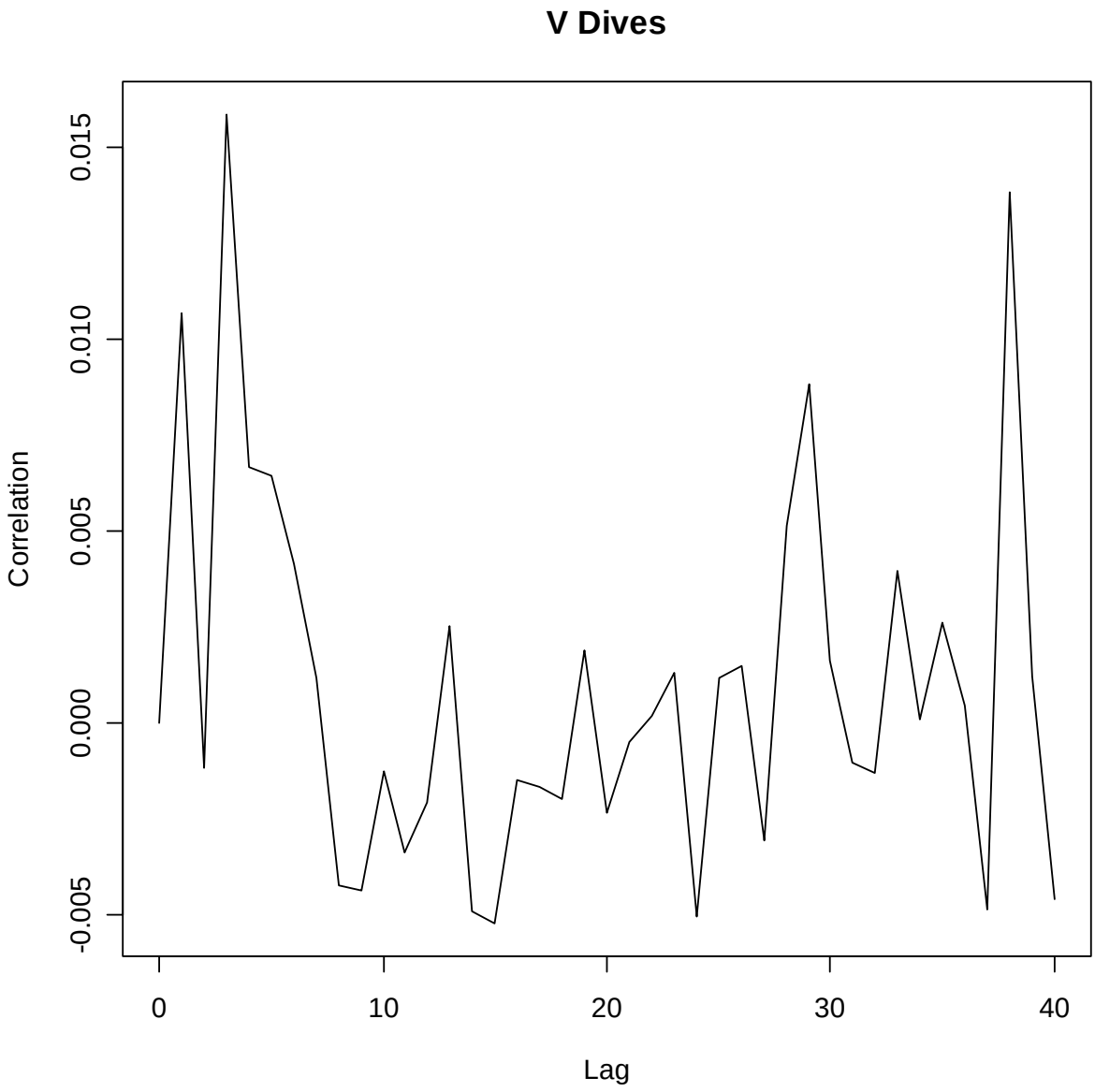

Figure 39. Correlogram of Pearson's residuals versus lag in dives for the GAM fitted to binary data for V dives.

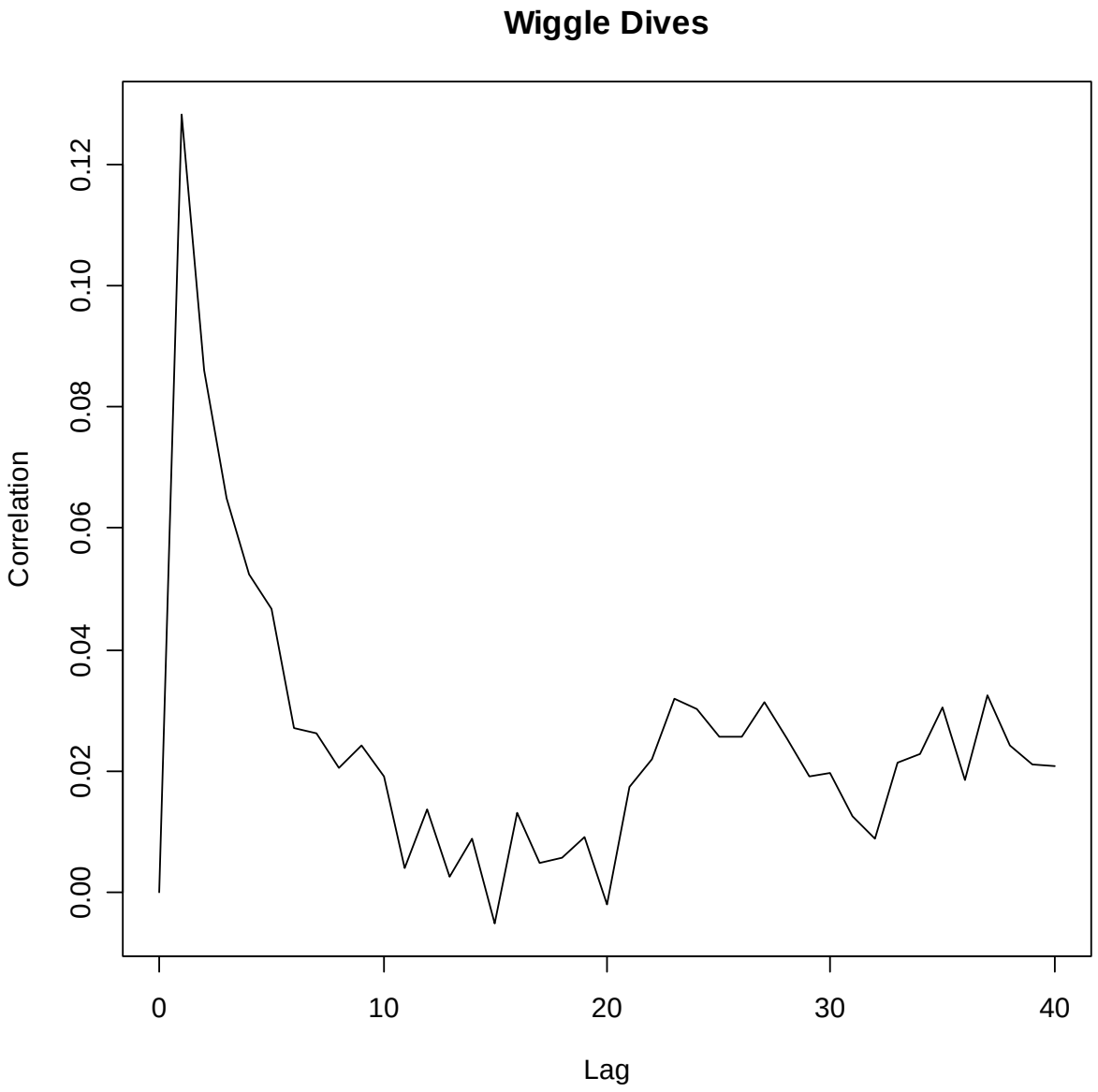

Figure 40. Correlogram of Pearson's residuals versus lag in dives for the GAM fitted to binary data for W dives.

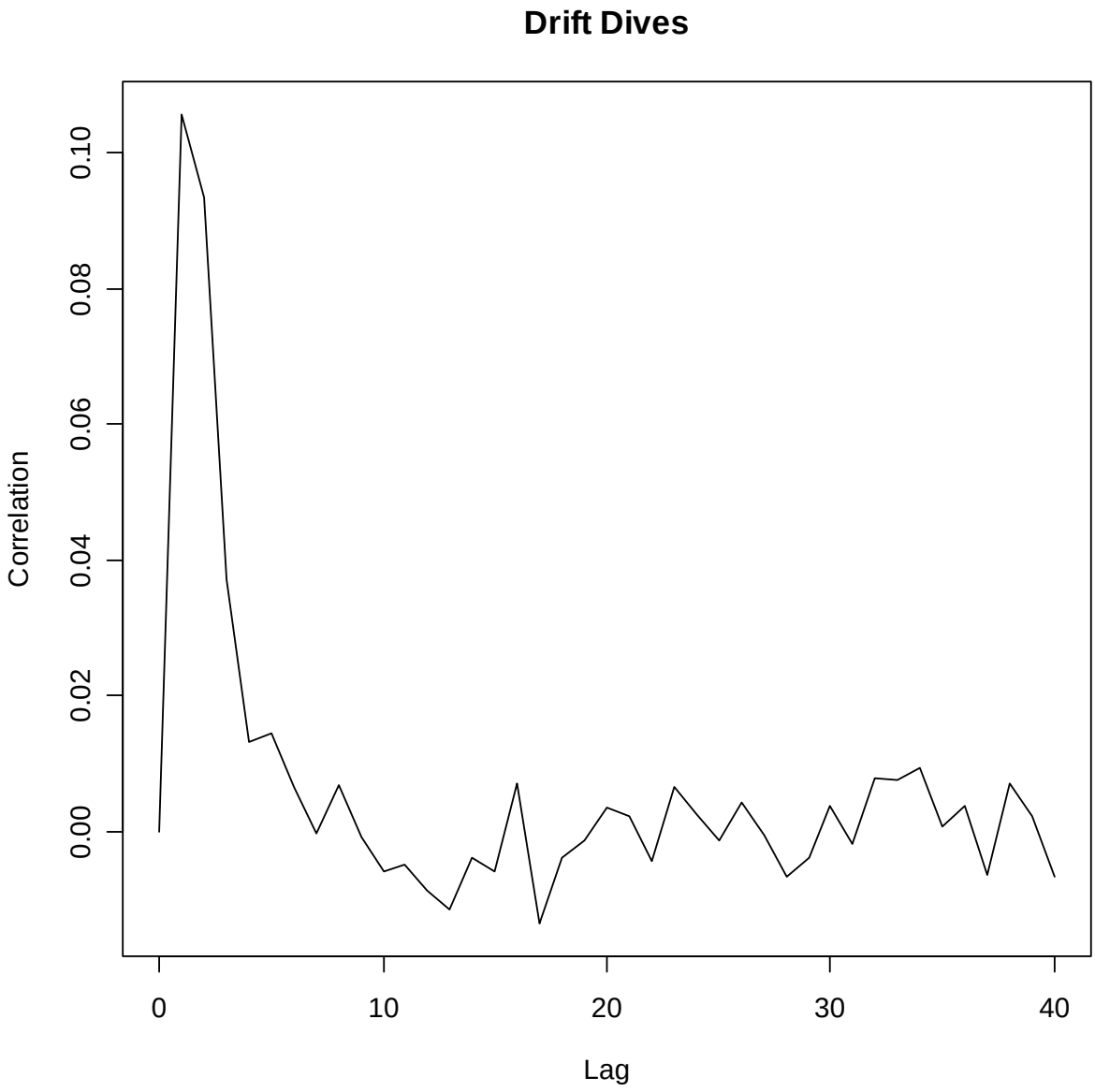

Figure 41. Correlogram of Pearson's residuals versus lag in dives for the GAM fitted to binary data for DR dives.

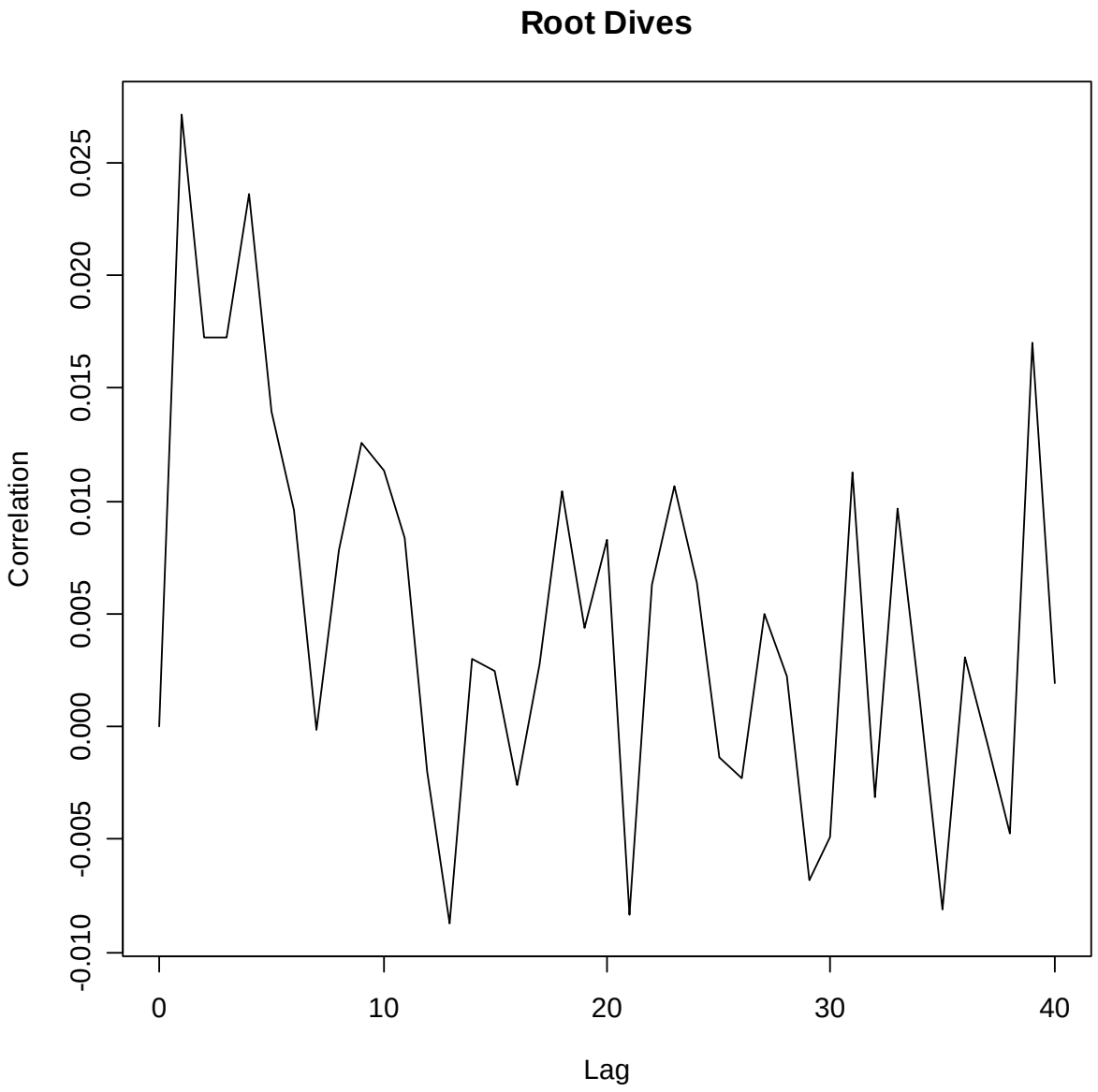

Figure 42. Correlogram of Pearson's residuals versus lag in dives for the GAM fitted to binary data for R dives.

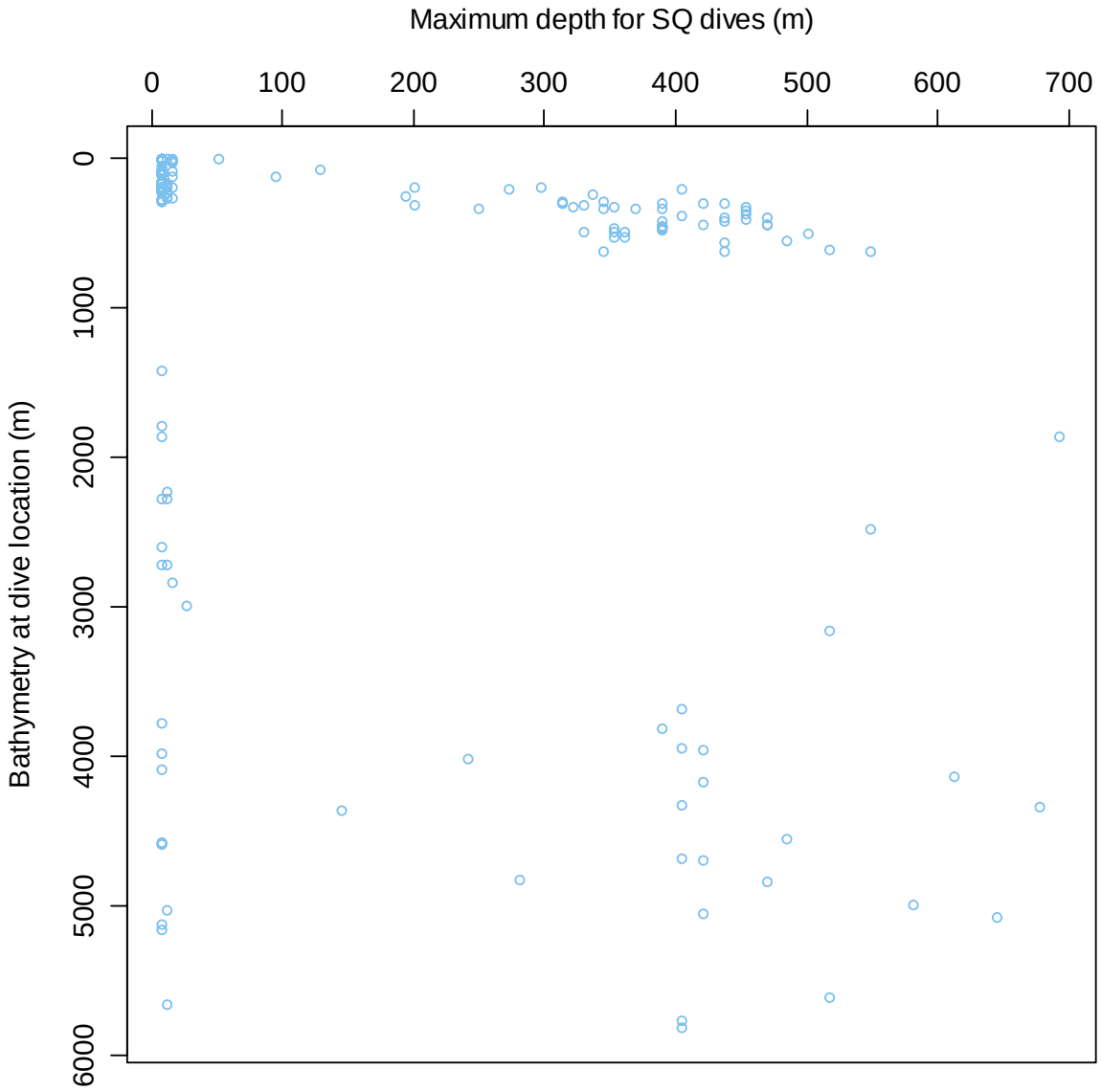

Figure 43. Relationship between maximum depth (m) and bathymetry (m) in SQ dives

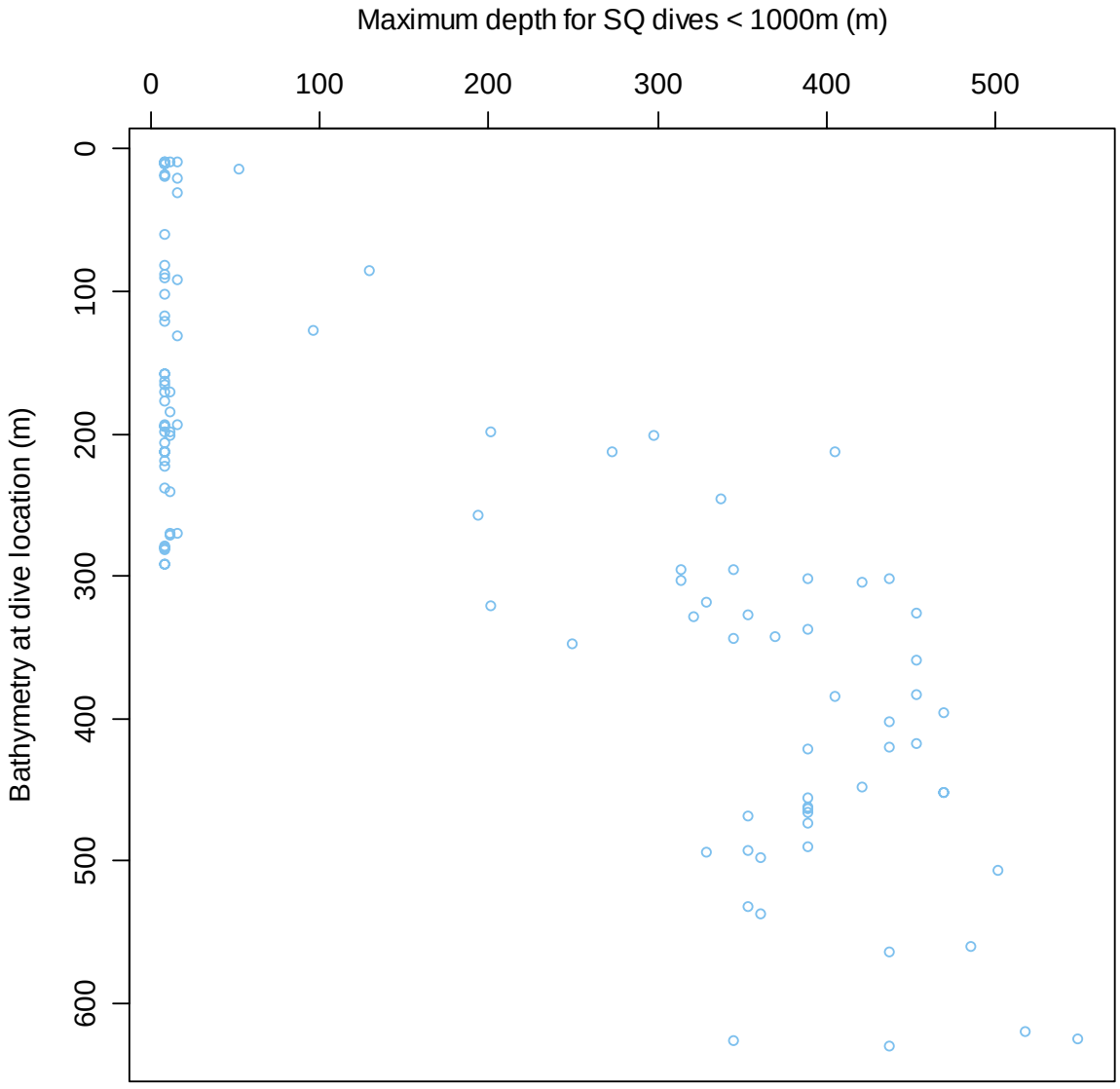

Figure 44. Relationship between maximum depth (m) and bathymetry (m) in SQ dives for dives shallower than 1000m

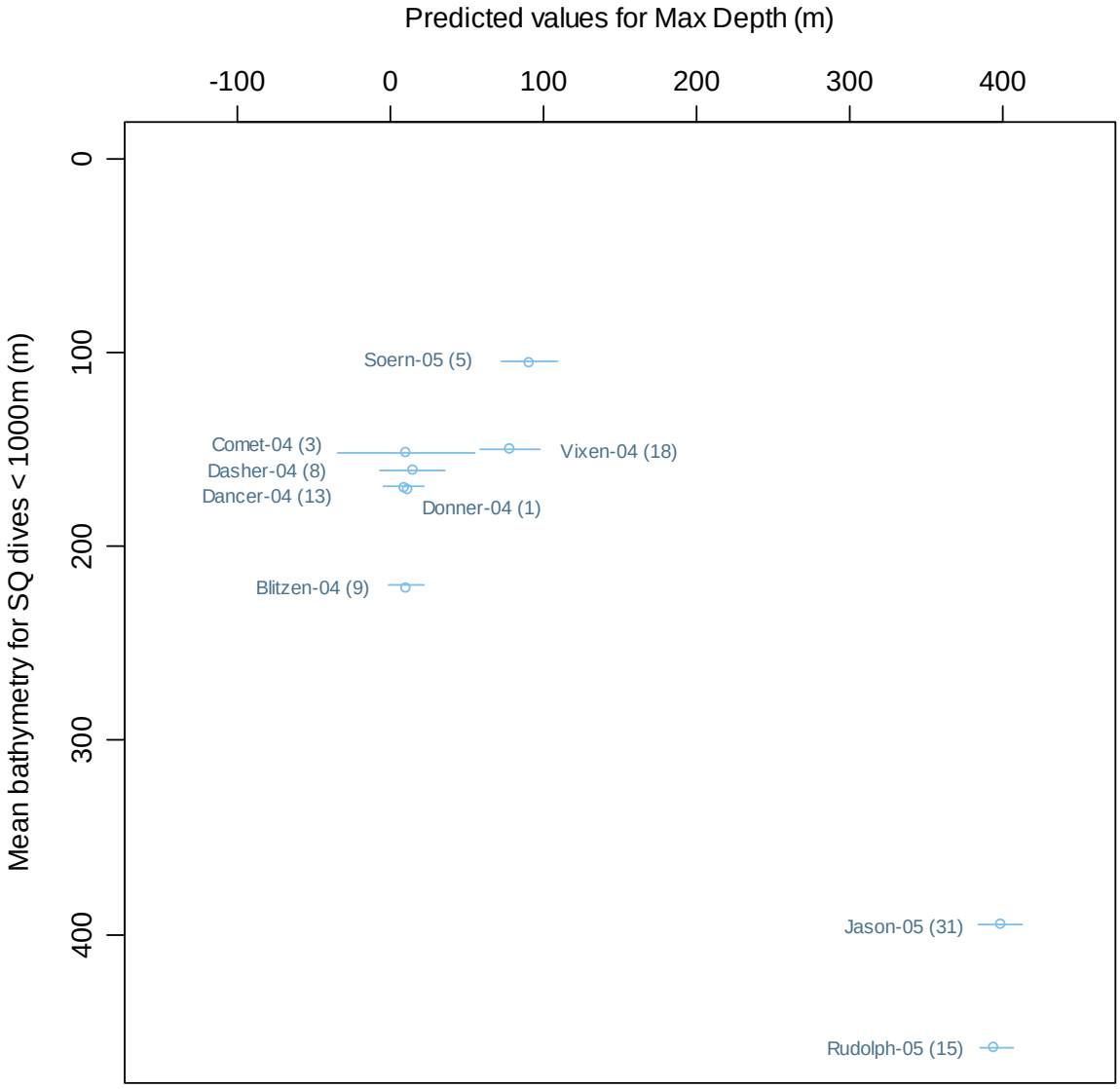

Figure 45. Individual predictions for mean maximum dive depth according to the GLS model fitted to the relationship between maximum dive depth and bathymetry in SQ dives.

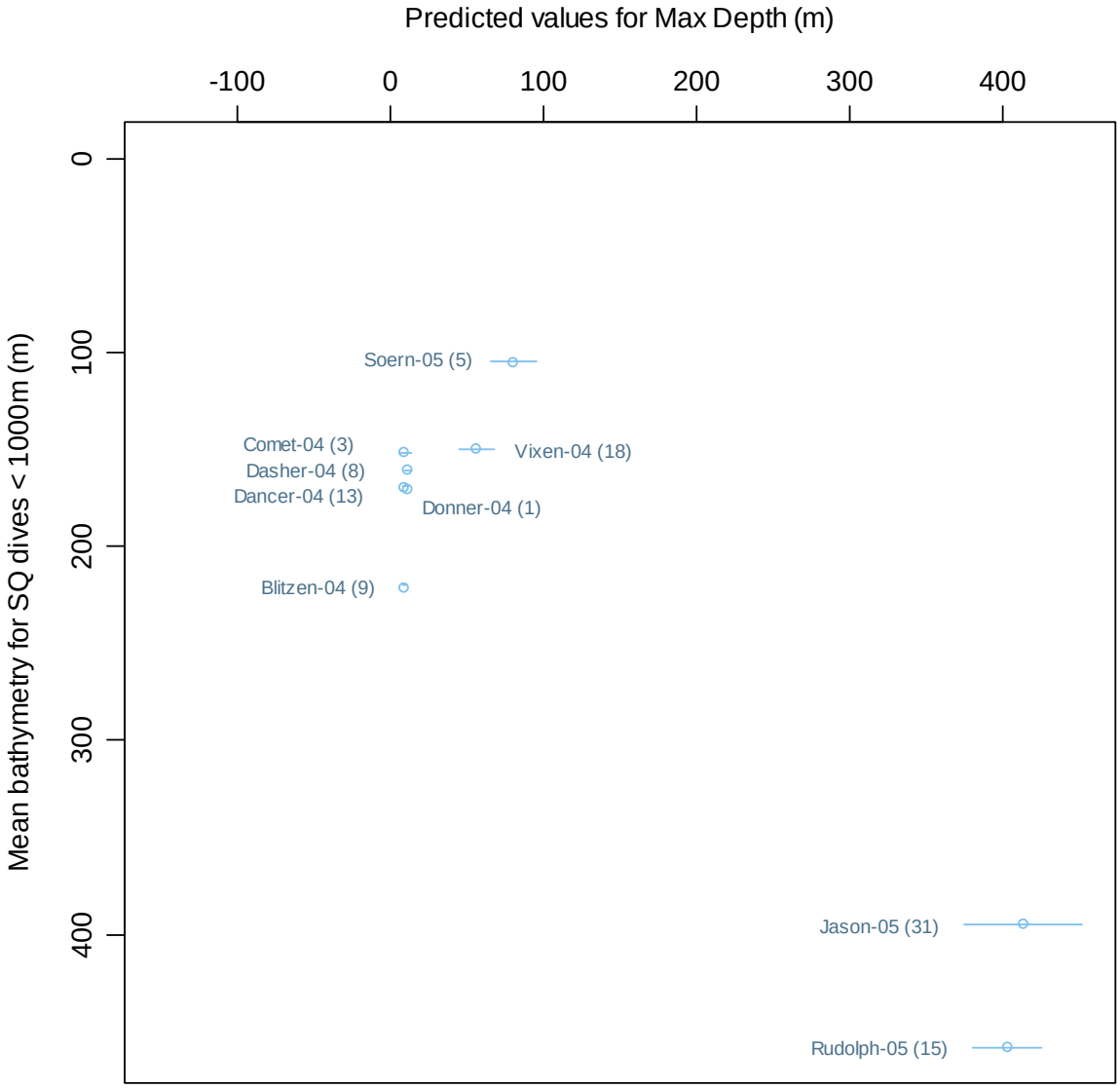

Figure 46. Individual predictions for mean maximum dive depth according to the GAMM model fitted to the relationship between maximum dive depth and bathymetry in SQ dives.

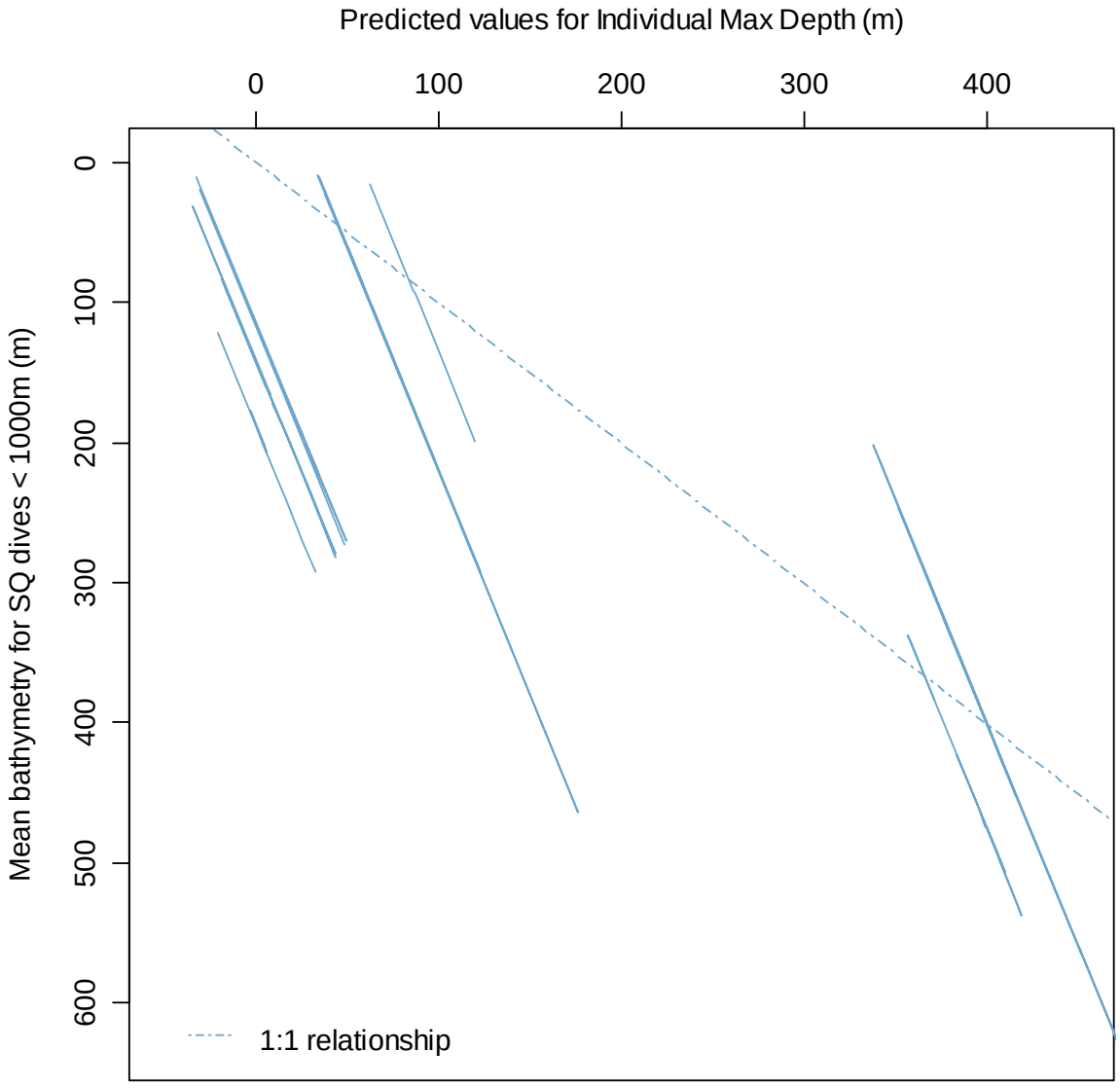

Figure 47. Individual predictions for the GLS model fitted to the relationship between maximum dive depth and bathymetry in SQ dives.

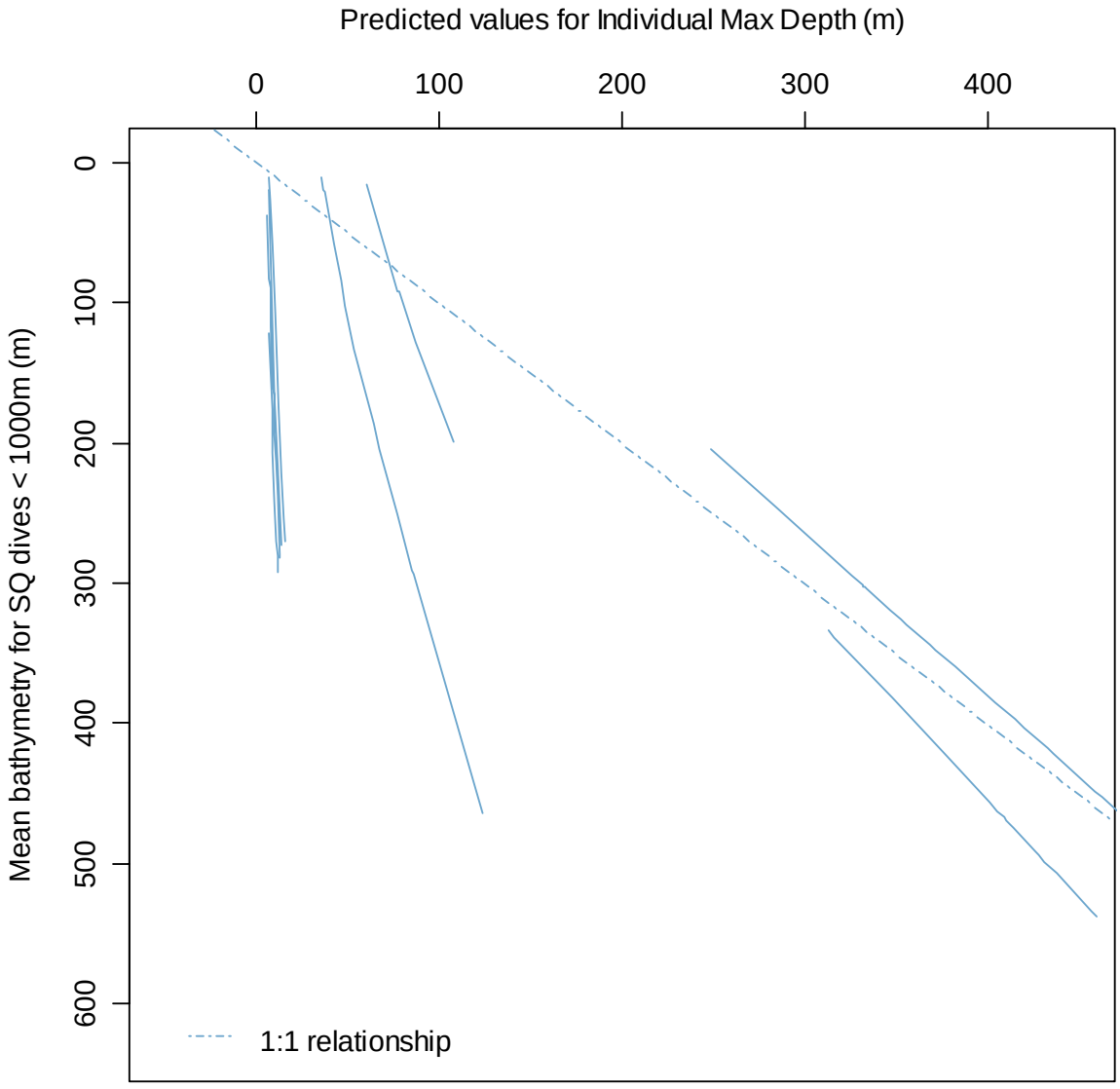

Figure 48. Individual predictions for the GAMM model fitted to the relationship between maximum dive depth and bathymetry in SQ dives.

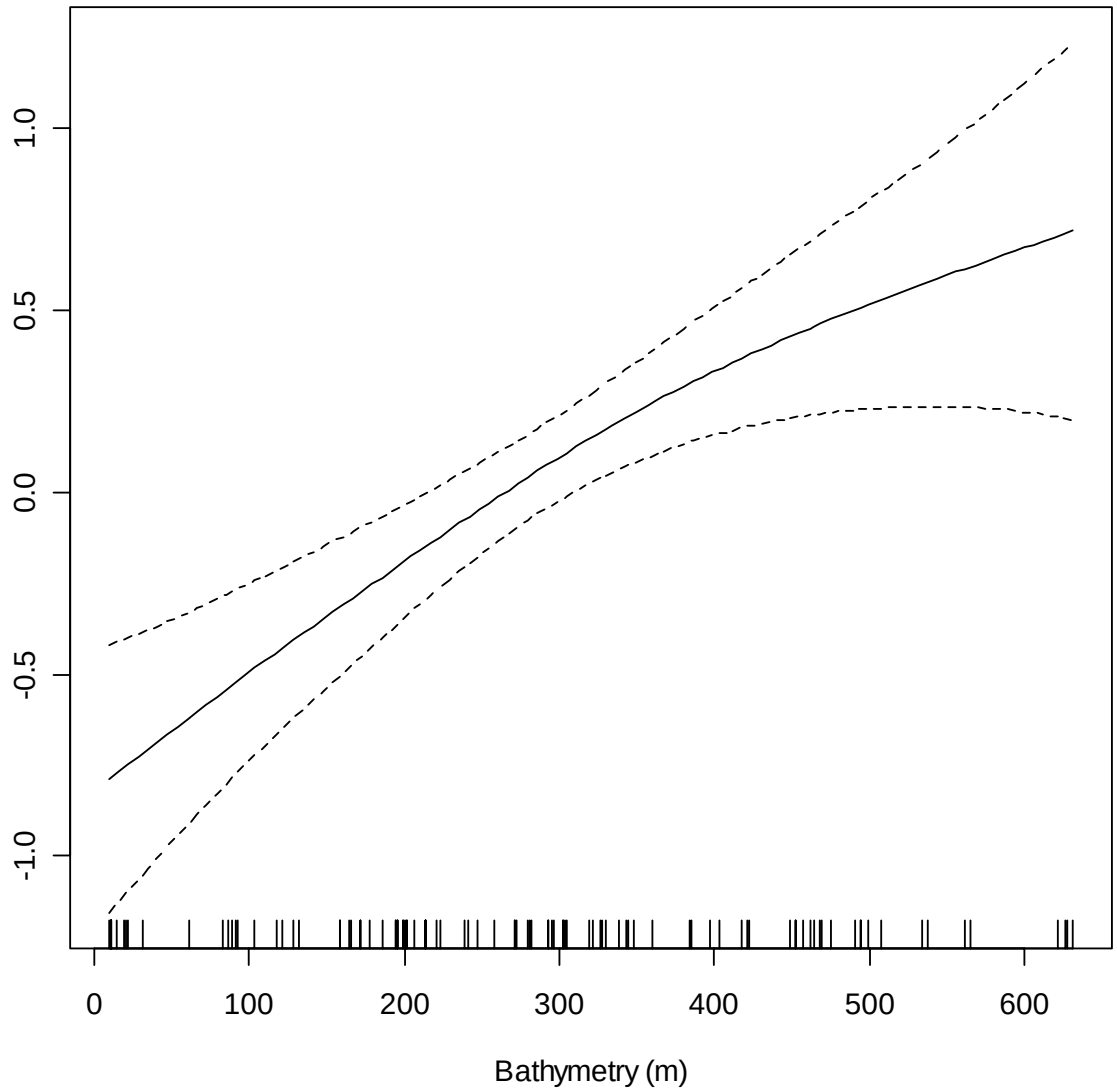

Figure 49. Smooth function for the GAMM model fitted to the relationship between maximum dive depth and bathymetry in SQ dives.